

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 21:37:39 ; Search time 1006 Seconds
(without alignments)
6908.530 Million cell updates/sec

Title: US-10-621-113-3

Perfect score: 1132

Sequence: 1 agccatggcagggcccccat.....agcagtgcgagggatccc 1132

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5700845 seqs, 3069779757 residues

Total number of hits satisfying chosen parameters: 11401690

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1132	100.0	1132	17 US-10-621-113-3	Sequence 3, Appli
2	1116	98.6	1129	17 US-10-621-113-1	Sequence 1, Appli
3	1107	97.8	1147	17 US-10-621-113-5	Sequence 5, Appli
4	1091	96.4	1147	17 US-10-621-113-7	Sequence 7, Appli
c	342.4	30.2	2922	17 US-10-108-260A-1465	Sequence 1465, Ap
5	60	5.3	1180	18 US-10-437-963-24598	Sequence 24598, A
6	60	5.3	2076	15 US-10-156-761-7099	Sequence 7099, Ap
7	58.2	5.1	9025608	15 US-10-156-761-1	Sequence 1, Appli
c	58.2	5.1	1041	18 US-10-437-963-8597	Sequence 8597, Ap
9	57.2	5.1	2781	18 US-10-357-930-23327	Sequence 23327, A
10	56.8	5.0	2781	18 US-10-357-930-23327	Sequence 23327, A
11	56.8	5.0	2781	18 US-10-357-930-23327	Sequence 23327, A

c	12	56.4	5.0	1521	18	US-10-425-115-94529	Sequence 94529, A
	13	55.6	4.9	2947	19	US-10-278-698-276	Sequence 276, App
	14	55.6	4.9	2947	19	US-10-278-698-790	Sequence 790, App
c	15	54.8	4.8	1240	17	US-10-425-114-31082	Sequence 31082, A
	16	54.8	4.8	4667	18	US-10-723-860-5759	Sequence 5759, Ap
	17	54.8	4.8	84428	17	US-10-229-1488-1	Sequence 1, Appli
	18	54.6	4.8	4056	17	US-10-388-934-653	Sequence 653, App
c	20	54	4.8	1041	18	US-10-437-963-8597	Sequence 8597, Ap
	21	54	4.8	1566	18	US-10-437-963-31282	Sequence 31282, A
c	22	53.2	4.7	1574	17	US-10-324-316-11	Sequence 11, Appli
	23	53.2	4.7	4545	17	US-10-347-470A-29	Sequence 29, Appli
	24	53.2	4.7	4545	18	US-10-322-281-585	Sequence 585, App
	25	53	4.7	95914	18	US-10-322-281-584	Sequence 584, App
	26	53	4.7	1575	15	US-10-156-761-7467	Sequence 7467, Ap
	27	52.8	4.7	9025608	15	US-10-156-761-1	Sequence 1, Appli
c	28	52.2	4.6	954	15	US-10-156-761-3934	Sequence 3934, Ap
	29	52	4.6	972	18	US-10-425-115-17653	Sequence 17653, A
	30	51.8	4.6	788	18	US-10-425-115-63819	Sequence 63819, A
	31	51.8	4.6	805	17	US-10-424-599-61903	Sequence 61903, A
	32	51.8	4.6	2156	9	US-09-795-651-61	Sequence 61, Appli
	33	51.8	4.6	2156	19	US-10-952-459-23	Sequence 23, Appli
	34	51.8	4.6	2359	18	US-10-473-974-171	Sequence 171, App
	35	51.6	4.6	2700	16	US-10-240-965-91	Sequence 91, Appli
c	36	51.6	4.6	763	18	US-10-437-963-49370	Sequence 49370, A
	37	51.6	4.6	1215	17	US-10-425-114-18849	Sequence 18849, A
c	38	51.6	4.6	2327	18	US-10-437-963-66835	Sequence 66835, A
	39	51.6	4.6	4667	18	US-10-723-860-5759	Sequence 5759, Ap
c	40	51.6	4.6	174424	9	US-09-967-768A-314	Sequence 314, App
	41	51.6	4.6	174424	10	US-09-960-706-969	Sequence 969, App
c	42	51.6	4.6	174424	19	US-10-843-641A-6459	Sequence 6459, Ap
	43	51.2	4.5	181343	18	US-10-723-860-2392	Sequence 2392, Ap
c	44	50.8	4.5	1373	18	US-10-437-963-83745	Sequence 83745, A
	45	50.6	4.5	1180	18	US-10-437-963-24598	Sequence 24598, A
c	45	50.6	4.5	1954	18	US-10-767-701-14885	Sequence 14885, A

ALIGNMENTS

RESULT 1
US-10-621-113-3
; Sequence 3, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1117)
; OTHER INFORMATION:
US-10-621-113-3

Query Match 100.0%; Score 1132; DB 17; Length 1132;
Best Local Similarity 100.0%; Pred. NQ. 2.5e-280;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCCATGGCAGGCCCGGATACCCAGTTTCAGTGCAGGGGAGCCCTGGTGCAGATCAA 60
DB 1 AGCCATGGCAGGCCCGGATACCCAGTTTCAGTGCAGGGGAGCCCTGGTGCAGATCAA 60

```
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGGACACCTTCGTGCG 120
DB |||||
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGGACACCTTCGTGCG 120
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAAGACCTCTCAAGAGACCTTCCTCCGGTGA 180
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAAGACCTCTCAAGAGACCTTCCTCCGGTGA 180
DB |||||
QY 181 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 240
DB |||||
QY 181 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 240
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCCGGCTGCGCGCTGAGCTGTTGGAACTATTC 300
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCCGGCTGCGCGCTGAGCTGTTGGAACTATTC 300
DB |||||
QY 301 TCGGAGGCTGTGGCGACTGACAGAGCGGTGGCAGCGCCGACGATCACTGGCTTCT 360
DB |||||
QY 301 TCGGAGGCTGTGGCGACTGACAGAGCGGTGGCAGCGCCGACGATCACTGGCTTCT 360
DB |||||
QY 361 CGCACCGCAACCTCTGGACCTGGAGCCCGCTGCCACCCCGCAGCCGGTGATCCTGCC 420
DB |||||
QY 361 CGCACCGCAACCTCTGGACCTGGAGCCCGCTGCCACCCCGCAGCCGGTGATCCTGCC 420
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTCGGCTGCGGGCCGCTTCCATCCACAGTCTGA 480
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTCGGCTGCGGGCCGCTTCCATCCACAGTCTGA 480
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGACGCTTCTGTATCCAGAGACGCGGATAGGCCCTT 540
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGACGCTTCTGTATCCAGAGACGCGGATAGGCCCTT 540
DB |||||
QY 541 TCAGGCGCAGGCCCAGGAGAGCTTGACGTGCTGCGGACCCCTCAGGCTGGTGGCT 600
DB |||||
QY 541 TCAGGCGCAGGCCCAGGAGAGCTTGACGTGCTGCGGACCCCTCAGGCTGGTGGCT 600
DB |||||
QY 601 GGTGAGAACGAAGACCGGACACCGCTGTTTCAGGAGCAGCGTCCCGAGTCTGTGC 720
DB |||||
QY 601 GGTGAGAACGAAGACCGGACACCGCTGTTTCAGGAGCAGCGTCCCGAGTCTGTGC 720
DB |||||
QY 721 TTCCGCGCTTACGAGAGCAGCCGCGCAGATGAGCTGTCGTGCCCGGGGGCGCGT 780
DB |||||
QY 721 TTCCGCGCTTACGAGAGCAGCCGCGCAGATGAGCTGTCGTGCCCGGGGGCGCGT 780
DB |||||
QY 781 GCGGTGTTGGAAACGTCAGACCGCGCTGTGGCTATGACAGGTACGGGACCGGGCGG 840
DB |||||
QY 781 GCGGTGTTGGAAACGTCAGACCGCGCTGTGGCTATGACAGGTACGGGACCGGGCGG 840
DB |||||
QY 841 CCTACTCCCGGGTGTGCTGCGGCGGAGAGGCTGGCGCTCTCTGAGCGGACCGG 900
DB |||||
QY 841 CCTACTCCCGGGTGTGCTGCGGCGGAGAGGCTGGCGCTCTCTGAGCGGACCGG 900
DB |||||
QY 901 GTTTCGTGGAGAGACGACCCGCGGGGTGAGGCCCGGGGCTTCCCTGAACCTCCAGGC 960
DB |||||
QY 901 GTTTCGTGGAGAGACGACCCGCGGGGTGAGGCCCGGGGCTTCCCTGAACCTCCAGGC 960
DB |||||
QY 961 CACCGCCCTCCCGCAGCGTCCACCGACCTTTCGCGGGGCGCATCCAGAGCGCTG 1020
DB |||||
QY 961 CACCGCCCTCCCGCAGCGTCCACCGACCTTTCGCGGGGCGCATCCAGAGCGCTG 1020
DB |||||
QY 1021 CTGACCGCTCAGACGAGGGCTCTGAGCGGGCCACGCGGCGGAGGCGCGCTTCGAGG 1080
DB |||||
QY 1021 CTGACCGCTCAGACGAGGGCTCTGAGCGGGCCACGCGGCGGAGGCGCGCTTCGAGG 1080
DB |||||
QY 1081 GTGCGTGAATCTGTGCGCGACCCACGACGAGCAGTGAAGCGGAGATCC 1132
DB |||||
QY 1081 GTGCGTGAATCTGTGCGCGACCCACGACGAGCAGTGAAGCGGAGATCC 1132
DB |||||
```

```
RESULT 2
US-10-621-113-1
; Sequence 1, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621.113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1129
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1114)
; OTHER INFORMATION:
US-10-621-113-1
```

Query Match 98.6%; Score 1116; DB 17; Length 1129;
Best Local Similarity 99.7%; Pred. No. 3.1e-276;
Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
DB |||||
QY 1 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
DB |||||
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGGACACCTTCGTGCG 120
DB |||||
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGGACACCTTCGTGCG 120
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAAGACCTCAAGGAGACCTTCCCGGTGA 180
DB |||||
QY 121 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAAGACCTCAAGGAGACCTTCCCGGTGA 177
DB |||||
QY 181 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 240
DB |||||
QY 178 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTCTCGATGACCACTGTT 237
DB |||||
QY 241 GGGACGCTGGGGCCGACGAGCCCGCTGCGCGCTGCGAGCTTGGAAACCTATTC 300
DB |||||
QY 238 GGGACGCTGGGGCCGACGAGCCCGCTGCGCGCTGCGAGCTTGGAAACCTATTC 297
DB |||||
QY 301 TCGGAGGCTGTGCGAGCTGCGAGCGCTGCGAGCGCCGAGCCCGAGCTTCTTCTT 360
DB |||||
QY 298 TCGGAGGCTGTGCGAGCTGCGAGCGCTGCGAGCGCCGAGCCCGAGCTTCTTCTT 357
DB |||||
QY 361 GCGACCGCAACCTCTGAGCTGAGCGCGCTGCGACCCCGGAGCCGGGTGATCCTGCC 420
DB |||||
QY 358 GCGACCGCAACCTCTGAGCTGAGCGCGCTGCGACCCCGGAGCCGGGTGATCCTGCC 417
DB |||||
QY 421 CACCCAGAGGAGCAGCTTCTTCTCGGCTGCGGGCGCTTCCATCCACAGTCTGA 480
DB |||||
QY 418 CACCCAGAGGAGCAGCTTCTTCTCGGCTGCGGGCGCTTCTCCATCCACAGTCTGA 477
DB |||||
QY 481 GGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGACACGCGGGATAGGCCCTT 540
DB |||||
QY 478 GGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGACACGCGGGATAGGCCCTT 537
DB |||||
QY 541 TCAGGCGCAGGCCCAGAGAGCTTGGAGCTGCTGCGGCGAGCCCTCAGGCTGGTGGCT 600
DB |||||
QY 538 TCAGGCGCAGGCCCAGAGAGCTTGGAGCTGCTGCGGCGAGCCCTCAGGCTGGTGGCT 597
DB |||||
QY 601 GGTGAGAACGAAGACCGGACCGCTGTGTTTTCCAGCGCCCTACCTGGAGGAGGCGCG 660
DB |||||
```

```

Db      598  |||||GGTGGAGAAAGAAAGACCGGAGAGACCGGCTGTGTTTCCAGCGCCCTACCTGGAGAGGCGGC 657
Qy      661  |||||CCCGGGCCAAAGGCGGAGAGGAGCCCGTCCCTAGGAGAGAGCGGTCCCCAGTCTCTGC 720
Db      658  |||||CCCGGGCCAAAGGCGGAGAGGAGCCCGTCCCTAGGAGAGAGCGGTCCCCAGTCTCTGC 717
Qy      721  |||||TTCCCGGCGCTACGAGAGAGCGCGCAGATGAGCTCTCGTGCCTCCCGGGGGCGCGCGT 780
Db      718  |||||TTCCCGGCGCTACGAGAGAGCGCGCAGATGAGCTCTCGTGCCTCCCGGGGGCGCGCGT 777
Qy      781  |||||GCCGGTGTGGAAACGTCAGACCGCGGCTGGTGGCTATGAGGTACGGCCACCGGGGGG 840
Db      778  |||||GCCGGTGTGGAAACGTCAGACCGCGGCTGGTGGCTATGAGGTACGGCCACCGGGGGG 837
Qy      841  |||||CCTACTCCCCCGGCTGTGTCGGCCGGAAGGCTGGGCGCTCTCTGAGCGGAGCGGG 900
Db      838  |||||CCTACTCCCCCGGCTGTGTCGGCCGGAAGGCTGGGCGCTCTCTGAGCGGAGCGGG 897
Qy      901  |||||GTTCCGTGGAGAGACAGACCCGCGGTGAGGCCCGGGCTTCCCTGAACCTTCCAGGC 960
Db      898  |||||GTTCCGTGGAGAGACAGACCCGCGGTGAGGCCCGGGCTTCCCTGAACCTTCCAGGC 957
Qy      961  |||||CACGGCCCTCCCCCACCCTGCGCCACCGCTTGGCCCGGCGGCGCATCCAGAGCCGCTG 1020
Db      958  |||||CACGGCCCTCCCCCACCCTGCGCCACCGCTTGGCCCGGCGGCGCATCCAGAGCCGCTG 1017
Qy     1021  |||||CTGCACCGTCACACGAGGAGCCCTGGAGCGGCGGCCACCGCGCCAGGCGCGCCCTCGAGG 1080
Db     1018  |||||CTGCACCGTCACACGAGGAGCCCTGGAGCGGCGGCCACCGCGCGCAGGCGCGCCCTCGAGG 1077
Qy     1081  |||||GTGCGTGGACTCTGTGCGCCAGCCACCGAGCGGAGCAGTGAAGCGGAGGATCC 1132
Db     1078  |||||GTGCGTGGACTCTGTGCGCCAGCCACCGAGCGGAGCAGTGAAGCGGAGGATCC 1129

```

RESULT 3

```

US-10-621-113-5
; Sequence 5, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621.113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1147
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1132)
; OTHER INFORMATION:
US-10-621-113-5

```

```

Query Match      97.8%; Score 1107; DB 17; Length 1147;
Best Local Similarity 98.7%; Pred. No. 6.3e-274;
Matches 1132; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Qy      1  AGCCATGGCAGGCCCCCGATACCAGTTTCAGTGCAAGGGGACGCGCTGGTGCAGATCAA 60
Db      1  AGCCATGGCAGGCCCCCGATACCAGTTTCAGTGCAAGGGGACGCGCTGGTGCAGATCAA 60
Qy     61  GAGGCTCCAAACGTTTGCCTTCTCTGTGGCGTGGTCAAGCGGAGCACACCTTCGTGGC 120

```

```

Db      61  GAGGCTCCAAACGTTTGCCTTCTCTGTGGCGTGGTCAAGCGGAGCGACACCTTCGTGGC 120
Qy     121  CAGGAGTTGGAGCAAAATTCAGGCGAGCTCAAGAAAGACCTTCAAGGAGACCTTCCCGGTGGA 180
Db     121  CAGGAGTTGGAGCAAAATTCAGGCGAGCTCAAGAAAGACCTTCAAGGAGACCTTCCCGGTGGA 180
Qy     181  GCGGCGCTGCTCGGAGATCTGACCGCGTTCTCCAAAGCTTCTC----- 226
Db     181  GCGGCGCTGCTCGGAGATCTGACCGCGTTCTCCAAAGCTTCTC----- 240
Qy     227  -GATGCACACACTGTTTGGGAGCGGTGGGCGCAGACGCGGGCGCTGGCGGCTTCGAGCT 285
Db     241  GGATGCACCACTGTTTGGGAGCGGTGGGCGCAGACGCGGGCGCTGGCGGCTTCGAGCT 300
Qy     286  GTTGGAAAACCTATTTCTCGGAGGCTGTGGCGACTGACAGAGCGGTGGCAGGAGCCGAC 345
Db     301  GTTGGAAAACCTATTTCTCGGAGGCTGTGGCGACTGACAGAGCGGTGGCAGGAGCCGAC 360
Qy     346  GATCACTGGGTTCTTCGACACCGCAACCTTGACCTTGAGACCTGGAGCCGCGCTGCAACCCGCGCAG 405
Db     361  GATCACTGGGTTCTTCGACACCGCAACCTTGAGACTTGAGAGCCGCGCTGCAACCCGCGCAG 420
Qy     406  CCGGCTGATCTGCTGCCACCGCCAGAGAGAGAGCTCTTTCTCGCGCTGCGGGCGCGCTCTC 465
Db     421  CCGGCTGATCTGCTGCCACCGCCAGAGAGAGAGCTCTTTCTCGCGCTGCGGGCGCGCTCTC 480
Qy     466  CATCCACAGTCTGGAGGCTCAGAGCGCTGCGCTGCGCTGAGCCCTTCTGTACCCAGGACAC 525
Db     481  CATCCACAGTCTGGAGGCTCAGAGCGCTGCGCTGCGCTGAGCCCTTCTGTACCCAGGACAC 540
Qy     526  GCGGAGTAGGCTTTTTCAGCGCGCAGGCCAGGAGAGCTTGAGACCTGCTGCTGCGGCAACC 585
Db     541  GCGGAGTAGGCTTTTTCAGCGCGCAGGCCAGGAGAGCTTGAGACCTGCTGCTGCGGCAACC 600
Qy     586  CTCAGGCTGGTGGCTGGTGGAGAACGAAAGACCGGAGAGCGCTGGTGTTCAGGCGGCTTA 645
Db     601  CTCAGGCTGGTGGCTGGTGGAGAACGAAAGACCGGAGAGCGCTGGTGTTCAGGCGGCTTA 660
Qy     646  CTTGGAGGAGGCGGCGCCCGGCGCAAGCGCGGAGGAGGAGCCCTCTTCTAGGAGAGAGCGG 705
Db     661  CTTGGAGGAGGCGGCGCCCGGCGCAAGCGCGGAGGAGGAGCCCTCTTCTAGGAGAGAGCGG 720
Qy     706  TCCCCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCCGCGCAGATGAGCTGTCCGTGCC 765
Db     721  TCCCCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCCGCGCAGATGAGCTGTCCGTGCC 780
Qy     766  CCGCGGGGCGCGGCTGCGCGTGTGGAAAACGTCAGACCGCGGCTGGTGGCTATGCAAGTA 825
Db     781  CCGCGGGGCGCGGCTGCGCGTGTGGAAAACGTCAGACCGCGGCTGGTGGCTATGCAAGTA 840
Qy     826  CCGCGACACCGGCGGCGCTACTCTCCCGCGGTGCTGCTGCGCGGAGAGGCTGGGCGCTCT 885
Db     841  CCGCGACACCGGCGGCGCTACTCTCCCGCGGTGCTGCTGCGCGGAGAGGCTGGGCGCTCT 900
Qy     886  CTTGAGCGGAGCAGGGGTTCCGTGGAGAGACGACCCGCGGGGTGAGGCGCCCGGGGCTTCCC 945
Db     901  CTTGAGCGGAGCAGGGGTTCCGTGGAGAGACGACACCCGCGGGGTGAGGCGCCCGGGGCTTCCC 960
Qy     946  TGAACCCCTCCAGGCGCACCGCGCTTCCCGCGCGCTGCGCGGAGAGGCTTGGCGGGGCGC 1005
Db     961  TGAACCCCTCCAGGCGCACCGCGCTTCCCGCGCGCTGCGCGGAGAGGCTTGGCGGGGCGC 1020
Qy    1006  CATTCAGAGCGCTGCTGCACCGTCACAGCAGGCGGCGCTGGAGCGGCGGCCACGCGGCCA 1065
Db    1021  CATTCAGAGCGCTGCTGCACCGTCACAGCAGGCGGCGCTGGAGCGGCGGCCACGCGGCCA 1080
Qy    1066  GGGCGCGCTTCAAGGGTGGTGGACTCTGTGCGCACCCCGCACGAGCGGAGCAGTGAAGCGG 1125
Db    1081  GGGCGCGCTTCAAGGGTGGTGGACTCTGTGCGCGCACCCCGCACGAGCGGAGCAGTGAAGCGG 1140
Qy    1126  AGGATCC 1132
Db    1141  AGGATCC 1147

```

```
RESULT 4
US-10-621-113-7
; Sequence 7, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1144
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1129)
; OTHER INFORMATION:
US-10-621-113-7

Query Match      96.4%; Score 1091; DB 17; Length 1144;
Best Local Similarity 98.4%; Pred. No. 7.9e-270;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;

QY      1  AGCCATGGCAGGCCCCCGATACCCAGTTCAGTGAAGGGGAGCCCTGGTGCAGATCAA 60
DB      1  AGCCATGGCAGGCCCCCGATACCCAGTTCAGTGAAGGGGAGCCCTGGTGCAGATCAA 60

QY      61  GAGGCTCCAAAGCTTTGGCTTCTCTGTGCGTGTGTCAGACGGCAGGACACCTTCGTGG 120
DB      61  GAGGCTCCAAAGCTTTGGCTTCTCTGTGCGTGTGTCAGACGGCAGGACACCTTCGTGG 120

QY      121  CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAGACCTCAAGAGACCTTCCCGGTGGA 180
DB      121  CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAGACCTTCCCGGTGGA 177

QY      181  GCGGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCCGCTCAGGCCAGCT 226
DB      178  GCGGGGCTGTCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCCGCTCAGGCCAGCT 237

QY      227  GATCACCACCTGTTGGGACGGTGGGGGCGACGAGCCGCGCTGGCGCGCTGCAGCT 285
DB      238  GGATCACCACCTGTTGGGACGGTGGGGGCGACGAGCCGCGCTGGCGCGCTGCAGCT 297

QY      286  GTTGGAAACCTATTCTCGGAGCTGTGCGGACTGACGCGCTGCGAGCGCTGGCAGCGCCGAC 345
DB      298  GTTGGAAACCTATTCTCGGAGCTGTGCGGACTGACGCGCTGCGAGCGCTGGCAGCGCCGAC 357

QY      346  GATCACTGGCTTCTTGCACCCGCAACCCCTGACCTTGAGCGCCGCGCTGCCACCCGGCAG 405
DB      358  GATCACTGGCTTCTTGCACCCGCAACCCCTGACCTTGAGCGCCGCGCTGCCACCCGGCAG 417

QY      406  CCGGTGATCTTCCGACCCGACGAGCGCTTCTTCTGCGCTGCGGGCGCGCTCTC 465
DB      418  CCGGTGATCTTCCGACCCGACGAGCGCTTCTTCTGCGCTGCGGGCGCGCTCTC 477

QY      466  CATCCACAGCTTGGAGGCTCAGAGCTGCGCTGCTGCGAGCCCTTCTGTACCCAGGACAC 525
DB      478  CATCCACAGCTTGGAGGCTCAGAGCTGCGCTGCTGCGAGCCCTTCTGTACCCAGGACAC 537

QY      526  GCGGATAGGCTTTTCAGGCGCAGGCGCAGGAGAGCTTGACGTGCTGCTGCGGACACC 585
DB      538  GCGGATAGGCTTTTCAGGCGCAGGCGCAGGAGAGCTTGAGAGCTTGAGAGCTGCTGCGGACACC 597

586  CTCAGGCTGGTGGCTGGAGAACAAAGACCGGCAGACCGGCTGGTTCAGAGCGCCCTA 645
598  CTCAGGCTGGTGGCTGGTGGAGAACAAAGACCGGCAGACCGGCTGGTTCAGAGCGCCCTA 657
646  CTTGAGAGAGCGCGCCCGGGCCAAAGCCGGAGAGGAGCCCGTCCCTTAGGAGAGCAGCG 705
658  CTTGAGAGAGCGCGCCCGGGCCAAAGCCGGAGAGGAGCCCGTCCCTTAGGAGAGCAGCG 717
706  TCCCAGATTCTGTGCTTCCCGGCTTACGAGAGAGCGCGCGCAGATGAGCTGTCTGTCGC 765
718  TCCCAGATTCTGTGCTTCCCGGCTTACGAGAGAGCGCGCGCAGATGAGCTGTCTGTCGC 777
766  CGCGGGCGCGCGCTGCGCTGTGGAACCTGACACCGCGCTGCTGCTGCTATGAGGTA 825
778  CGCGGGCGCGCGCTGCGCTGTGGAACCTGACACCGCGCTGCTGCTGCTATGAGGTA 837
826  CGCGACCGCGCGGCTTACTTCCCGCGGTGCTGCTGCGCGGAGGAGGCTGGGCGCTCT 885
838  CGCGACCGCGCGGCTTACTTCCCGCGGTGCTGCTGCGCGGAGGAGGCTGGGCGCTCT 897
886  CTTGAGCGGAGCGGGCTTCCGTTGAGAGAGACGACCGCGCGGCTGAGGCGCGGGCTTCCC 945
898  CTTGAGCGGAGCGGGCTTCCGTTGAGAGAGACGACCGCGCGGCTGAGGCGCGGGCTTCCC 957
946  TGAACCTTCCAGAGCGCGCTTCTCCCGCCACCGTGCACCGTGCACCGTGCACCGCGCG 1005
958  TGAACCTTCCAGAGCGCGCTTCTCCCGCCACCGTGCACCGTGCACCGTGCACCGCGCG 1017
1006  CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
1018  CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
1066  GGGCGCGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
1078  GGGCGCGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1137
1126  AGGATCC 1132
1138  AGGATCC 1144

RESULT 5
US-10-108-260A-1465/c
; Sequence 1465, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1465
; LENGTH: 2922
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1465

Query Match      30.2%; Score 342.4; DB 17; Length 2922;
Best Local Similarity 84.9%; Pred. No. 3.9e-78;
Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;

QY      704  GGTCCCGAGTTCTGTGCTTCCCGCTTACGAGAGCAGCGCGCAGATGAGCTGTCCGTG 763
DB      679  GGTCCCGAGTTCTGTGCTTCCCGCTTACGAGAGCAGCGCGCAGATGAGCTGTCCGTG 620
764  CCCCGGGGGCGCGCTGCTGCTTGGAAACGTGACACCGCGCTGGTGGCTAT----- 818
619  CCCCGGGGGCGCGCTGCTGCTTGGAAACGTGACACCGCGCTGGTGGCTATGACAGG 560
819  ----- 818
```


Db 559 TACCGGGAGCGGGCGTGGGAGGGCAGGGTAGCCGAGGGGGGGGGCCCTTAACACAC 500
Qy 819 -----GCAGGTACGGCGACCGGGCGGGCCCTACTCCCGGGGTGCTGCTGCGGCCG 868
Db 499 CCGGCCCTCGCAGGTACGGGACCGGGCGGGCCCTACTCCCGGGGTGCTGCTGCGGCCG 440
Qy 869 GAAGGGTGGGGCTCTCTTGAAGCGGACCGGGTTCCTGGAGAGACGACCGGGCGGGT 928
Db 439 GAAGGGTGGGGCTCTCTTGAAGCGGACCGGGTTCCTGGAGAGACGACCGGGCGGGT 380
Qy 929 GAGCGCGGGGGCTTCCTCTGAACCTCCAGGCGACCGCCCTCCCGCCACCGTGCCAC 988
Db 379 GAGCGCGGGGGCTTCCTCTGAACCTCCAGGCGACCGCCCTCCCGCCACCGTGCCAC 320
Qy 989 CGACCTTCGCGGGCGGCATCCAGAGCGGCTGTGTGACCGCTCACACAGGGGCGCTGGAG 1048
Db 319 CGACCTTCGCGGGCGGCATCCAGAGCGGCTGTGTGACCGCTCACACAGGGGCGCTGGAG 260
Qy 1049 CGGGGCCACCGGCCAGGGCGCGCCCTCGAGGGTGCCTGGAAGTCTGTGCGGACCCCA 1108
Db 259 CGGGGCCACCGGCCAGGGCGCGCCCTCGAGGGTGCCTGGAAGTCTGTGCGGACCCCA 200
Qy 1109 ACGGAGCAGTGAGCGCGAGGATCC 1132
Db 199 ACGGAGCAGTGAGCGCGAGGATCC 176

RESULT 6
US-10-437-963-24598
; Sequence 24598, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24598
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1180)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29567C.1
US-10-437-963-24598

Query Match 5.3%; Score 60; DB 18; Length 1180;
Best Local Similarity 48.7%; Pred. No. 8.9e-06;
Matches 170; Conservative 0; Mismatches 178; Indels 1; Gaps 1;
Qy 759 CCGTGCCTCGGGGGCGCGGTGCGGTGTGGAACGTCAGACCGGGCTGTGCTAT 818
Db 308 CCGGGGGCGCGGGNCGCGCCCGCGCGGGGCGGGGCGCCCGCCCGCGGCGCG 367
Qy 819 GCAGGTACCGCGGCGGGGCTACTCCCGCGGTGCTGTGCGCGCGGAAGGCTGG 878
Db 368 CGGCGCGCGCGCCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 426
Qy 879 GCGCTCTCTGACGGGACCGGGGTTTCGTGAGAGACGACCGCGGGGTGAGCGCGGG 938
Db 427 GCCCGCGCGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 486

Qy 939 GCTTCTCTGAACCTCTCCAGGCGACCGGGCCCTCCCGCCACCGTCCCGACCTTCGC 998
Db 487 CCGCGCGGGCGCGGGCGCGCCCGCGCCCGCGCCCGCGCGCGCGCGCGCGCGCG 546
Qy 999 CGGGCGGCATTCAGAGCGGCTGTGTGACCGGTCAACAGAGGGGCGCTGGAGCGGCGCCAC 1058
Db 547 CCGCGCGCGCGCGCGCGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 606
Qy 1059 GCGCGCAGGGCGCGCCCTCGAGGGGTGCTGACTCTGTGCGCGCACCCAC 1107
Db 607 CCNCG 655

RESULT 7
US-10-156-761-7099
; Sequence 7099, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7099
; LENGTH: 2076
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2076)
US-10-156-761-7099

Query Match 5.1%; Score 58.2; DB 15; Length 2076;
Best Local Similarity 44.0%; Pred. No. 2.5e-05;
Matches 380; Conservative 0; Mismatches 478; Indels 5; Gaps 3;
Qy 263 CGCGGCTGCGCGCGCTGCAGCTGTTGGAAACCTATTCTCGGAGGCTGTGGCGACTGCA 322
Db 424 CGCGGCATCATGCACCGGGACGTGAAGCGCGGAACATCTGCTGCGCGCGGACCGCACG 483
Qy 323 GAGCGGTGGACGGAGCGCGGACGATCACTGGGTTCTTTCGACCGCAACCCCTGGACCTG 382
Db 484 GGGCGGCGGTACGCGCGCGTGTGCTCACCGACTACGGCATCTCCGTGTCAACCGGACGCG 543
Qy 383 GAGCGCGCTGCCACCGCGGAGCGGGTGATCTCTCCACACCGACAGAGAGAGCGCTCTT 442
Db 544 GGGAGACCGGTACACCTCTACGTCGGCGCTGTGTGCGGACCGCCCGGCTATCTGGGCGCG 603
Qy 443 TCTCGCGCTGCGGGCGCGCTCTCCATCCACAGTCTGAGGCTCAGAGCTCGGCTGCTG 502
Db 604 GAGCGGGACCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 663
Qy 503 CAGCCCTTGTGTACCGAGGACGCGGGGATAGGCTTTTTCAGGCGGAGGCGCCAGGAGAGC 562
Db 664 CTGTACTTCGGGGTTCGAGGGGTGCGGGCGCTTCGAGCGGGGACACGCTATCTCGCGAGGT 723
Qy 563 CTGGAGCTGCTGCGGCGACCCCTCAGGCTGCTGCTGTGAGAAACGAAGACCGGCGAG 622
Db 724 ACCGCGTGTGTGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 783
Qy 623 ACCGCGTGTGTTCCAGCGCGCTTACCTGAGAGGAGCGCGCGCGCGCGCGCGCGCGCG 682

```
Db 784 CTCGCGCGATGCTCGGGAAGACCCCGGCGAGCGGATCACGGCGGAGACACGAGGCG 843
Qy 683 GCGCGCTCTTCTAGGAGACAGCGTCCCCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGC 742
Db 844 GCGCTGTGCG--GCGATCGTACGCGCCAGCCCGCCATCCGCGCACCCAGTTGGA-CCTGGG 899
Qy 743 CCGCGAGATGACTGTCGTGCGCGCGGCGGCGCGCGTCCGCGTGTGGAACGTCAGAC 802
Db 900 GTGCGAGCGCGGTGGGCGGAGCGCGGTGGACCAACCGAGACCGGGTATCCGCGGCGACGG 959
Qy 803 CCGCGTGTGCTATGACAGTACGCGACGCGGCGGCGCTACTCCCGCGGTGCTGCTG 862
Db 960 GCGGCGGCGCGTCCGCGCGGGAACGGAACGTCGCGGCGGTACGAGACCGGCGGCGACCGG 1019
Qy 863 CCGCGGAGGCGTGGGCGCTCTCTGAGCGGGAACGCGGTTCCTGAGGAGACGACCGG 922
Db 1020 GTCCGCGGCGGAGCGCGCGCGGTTCACCGCCACCGCGGCGGACCGCGGGAACCGCGG 1079
Qy 923 GCGGCTGAGGCGCGGCGCTTCTGAAACCTCCAGGCGCACCGCGCTCCCGCCACCGTG 982
Db 1080 GCGGCGGCGACGCGCTCGG-AGCGGCGGCGCGGTGCGGCGGCGCGCGCGACCGTCCGCGC 1138
Qy 983 CCACCGCGACCTTTCGCGGCGCGCCATCCAGAGCGCGTGTGACCGTACACGCGAGGCGC 1042
Db 1139 AGGACCGGACGCGCGCGCGGTACGCGCCAGCAGACCGCGCGCACCCCTGTCAGCGCAGGAGC 1198
Qy 1043 CTGAGAGCGCGCCACCGCGCGGCGCGCGCGCTCCAGAGGTCGTCGAGACTCTGTGCGCGAC 1102
Db 1199 CCGGCGCGCCACGACCGCGGACCGCGCGCGCTCGCGGTACGCGCGCGCGCGCGCCACCG 1258
Qy 1103 CCCAGACGCGAGCAGTGGCGCG 1125
Db 1259 CCCCCCTCGTACGCGCGAGAGCGCG 1281
```

RESULT 8

US-10-156-761-1/c

; Sequence 1, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 1

; LENGTH: 9025608

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (4187715)

; OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Query Match

Best Local Similarity 5.1%; Score 58.2; DB 15; Length 9025608;

Matches 380; Conservative 44.0%; Pred. No. 1.4e-05;

Mismatches 478; Indels 5; Gaps 3;

Qy 263 CCGCGCTTGGCGCGCTGCGAGTGTGGAAACCTATTCTCGGAGGCTGTCGCGACTGCA 322

5

```
Db 8480488 CCGCGCATCATGCACCGCGGACGTGAAGCGCGGGAACATCTCTGCTGCGCGCGACCGCAGC 8480429
Qy 323 GAGCGCTGGACGAGAGCGCCGACGATCACTGCTTCTTTGGCAACCGCAACCCCTGGAACCTG 382
Db 8480428 GGGCGCGGTACGCGCGGTGCTGTCTCACCGACTACGCGCATCTCCGTGCAACCGGACGCG 8480369
Qy 383 GAGCCCGCTTCCACCGCGGAGCGGCTGATCTCTGCGCCACCGCGAGGAGCAGCCTCTT 442
Db 8480368 GCGAGAGACCCGGTACACCTTCACTGCGCGGTGCTGCGACCGCGCGGTATCTGCGCGCG 8480309
Qy 443 TCTCGCGCTGCGGCGCGCTCTCTCATCCACAGTCTCGAGAGGCTCAGAGCGCTGCGCTGCGCTG 502
Db 8480308 GAGCGGCGGACGCGGCGGCGCGCGACCGCGACCGCGCGGCTGCTCTGCTGGGCTGCGAC 8480249
Qy 503 CAGCCCTTCTGTATCCAGGACACGCGGGAATAGGCTTTTTCAGGCGCAGCGCCAGGAGAGC 562
Db 8480248 CTGTACTTTCGGGCTCGAGGGGTGCGGCGCTTTCGAGCGGGAACACGCACTCTCGCCGAGGTC 8480189
Qy 563 CTGAGAGTGTCTGCGGCGCGCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 622
Db 8480188 ACCGCGGTGCTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8480129
Qy 623 ACCGCTGCTTTCAGCGCGCTTCTCTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 682
Db 8480128 CTCGCGCGATGCTCTCGGAGGACCGCGGCGAGCGGATACCGCGGAGGACACCGGAGGCG 8480069
Qy 683 GCGCGCTCTTTCAGGAGAGCGGCTTCTCAGTCTTCTGCTTCTTCTGCGCGCTTCTGAGAGCAGC 742
Db 8480068 GCGCTGTGCT--GCGATCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8480013
Qy 743 CGCGCAGATGAGTGTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 802
Db 8480012 GTGCGAGCGCGCGTGGGCGGAGCGCGGTGGAACACCGAGACCGGGTATTCGCGGCGCA 8479953
Qy 803 CGCGCTGTGCTATGACAGTACGCGGACCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 862
Db 8479952 GCGCGGCGCGGTCCCGCGGCGGAAACGTAACGTCGCGCGCGGTACGGAACCGGCGGCA 8479893
Qy 863 CCGCGGAGAGGCTGGGCGCTCTCTGAGCGGAGACGCGGCTTCCGTGAGGAGGAGACGACCGG 922
Db 8479892 GTCCGCGGCGGAGCGCGCGCGGCTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8479833
Qy 923 CCGGCTGAGGCGCGCGGCTTCTCTGAAACCTTCCAGGCGCGCGCGCGCGCGCGCGCGCGCG 982
Db 8479832 GCGGCGCACGCGCTCTCGG-AGCGGCGCGCGGTGCGGCGCGCGCGCGCGCGCGCGCGCGCG 8479774
Qy 983 CCCACCGCGACTTCCCGCGGCGCGCATCCAGAGCGCGCTCTCTGACCGCTCAGCAGCAGGCGCC 1042
Db 8479773 AGGACCGGAAACGCGCGCGCGGTACCGCCACGACCGCGCGCGCGCGCGCGCGCGCGCG 8479714
Qy 1043 CTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1102
Db 8479713 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8479654
Qy 1103 CCCACGAGCGAGCAGTGGAGCGCG 1125
Db 8479653 CCCCCCTGTACGCGCGAGAGCGCG 8479631
```

RESULT 9

US-10-437-963-8597/c

; Sequence 8597, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

RESULT 11
US-10-357-930-29208
; Sequence 29208, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276

```

```
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 29208
; LENGTH: 2781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 2779, 2780, 2781
; OTHER INFORMATION: n = A,T,C or G
US-10-930-29208

Query Match          5.0%; Score 56.8; DB 18; Length 2781;
Best Local Similarity 47.8%; Pred. No. 5.5e-05;
Matches 198; Conservative 0; Mismatches 212; Indels 4; Gaps 1;

QY 205 CGCGCTTCTCCCAAGCTTCTCGATGACACCACTGTTGGAGCGCGTGGGCGCACGAGCGG 264
Db 17 CGGGTTCTTCCCGGGCTACGGCGGGCTTGCCCGAGCGTTGGCTGCCGAGGCTCGGCGG 76

QY 265 CGGCTGCGCGCGCTGCACTGCTTTGGAAACCTATTCTCGAGGCTGCTGGGACTGCGAGA 324
Db 77 GAGCGTGGAGCGCGCGCGCTGCCCCAGGAC----CGCGCGCGCGCTTTGTCCCGCGCC 132

QY 325 GGGCTGGGACGAGGACCGGACGATCACTGGCTTCTTGGACCGGAAACCCCTGGAGCTGGA 384
Db 133 GCCACCGCGCGCTGCGCGCGCGCCATGAGCGCGCGCGCGCTGCGCGCGGGTCCCGCT 192

QY 385 GCCCGCTGCGCACCGCGCGGCTGATCTGCGCCACCCAGAGGAGCAGCCTCTTTTC 444
Db 193 TCCGCTGCTGCTGCGGGGCTTGGCGTGTGGCGCGGAGTGGAGCGGATGTCT 252

QY 445 TCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGAGGCTCAGAGCTCGCGTGCCTGCA 504
Db 253 CTGAGGCGCTGCTGCGGACGGACACCGGATGCCACTCATCAGAAGGACTGCTCGCT 312

QY 505 GCCCTTCTGTACCCAGGACCGCGGATAGGCTTTTTCAGGCGGAGGCCAGGAGGCT 564
Db 313 GCCATATGCTACGGAATCCAAAGATGCAAGGATGTCAGGATGTCAGGAGCAGTGTCTGCCACAGCA 372

QY 565 GGAAGTGTGCTGCGCGCACCCCTCAGGCTGGTGGCTGTGGAGAACGAAGACCG 618
Db 373 GCTGAGGAGCTGCACTGTGCCACGGGATCAGCTTGCCCAACGAGGAGACCG 426

RESULT 12
US-10-425-115-94529/c
; Sequence 94529, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Placenta
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
```

```
; SEQ ID NO 94529
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_186203C.1
US-10-425-115-94529

Query Match          5.0%; Score 56.4; DB 18; Length 1521;
Best Local Similarity 46.4%; Pred. No. 7.3e-05;
Matches 218; Conservative 0; Mismatches 251; Indels 1; Gaps 1;

QY 646 CCTGGAGGAGCGCGCCCGGGCCCAAGCGCGGAGGAGGCGCGCTCCCTAGGAGGAGCAGCGG 705
Db 1004 CCAGACGTAGCAGTCTGCGGCGCGGCTGAGACGGCGGCTGTAGTAGTCTCTCGGCGC 945

QY 706 TCCCGAGTTCTGTCTTCCCGCGCTACGAGAGCAGCGCGCGCAGATGATGATGTTCGCTGCC 765
Db 944 GGCCAGCGTCCGCGCTGTCTTCCCAACGAGGCGCGCTAGAGCAGACGCGCTCGGCGT 885

QY 766 CGCGGGGCGCGCGTGCCTGCTGTTGGAAAAGTCAAGCGCGGCTGTGGCTATGAGGTA 825
Db 884 CGCGGGGCTGCGGAGGATGCGCGCTCGCAGTACTCTCTGCGCGCTTTCGCCGCTCCCCGG 825

QY 826 CGCGAGCGCGCGCGCTTACTTCCCGCGGTGCTGTCGGCGCGGAAGGGCTGGGCGCTCT 885
Db 824 CCACCTCTTGAAGAACCTGGCTGTGTTGACGAGGAGGAGGAGTTGCGCGGTTCCGCT 765

QY 886 CCTGAGCGGAGCGGGGTTTCCGTGAGGAGACGACCCGCGGGTGTGAGGCCCGGGGCTTCCC 945
Db 764 CGATCATCTGCGCGTAGTGGCGCTTGGCGGCGCGGTTGCTGCCACCGTCCGCGTCCC 705

QY 946 TGAACCTTCCAGGCGACCGCGCTTCCCGCACCGTGTGCCACCGCTTCCCGGGGCGC 1005
Db 704 CGACCCACCGCTTCCCGCGCTGCTGCGCGACGCGCTTCCCGCGCGCGCTGCGCGCGA 645

QY 1006 CATCCAGAGCGCGTGTGACCGCTCACAGCAGGCGCGCTGAGGCGCGCCACGCGCGCA 1065
Db 644 CGCCTGTCTCACGAACGCGAGCGCGCGGCGGCGGACACGTCCAGACCGCTGTTCGTCA 585

QY 1066 G-GGCGCGCTTCGAGGTTGCTGAGACTTGTGTGCGCGACCCCGACCGAG 1114
Db 584 GCAGCGAGCGAGCGGCGAGCGGCGACGCTCCAGCGCGCGCGCGCGCGCGCGCG 535

RESULT 13
US-10-698-276
; Sequence 276, Application US/10278698
; Publication No. US20050037344A1
; GENERAL INFORMATION:
; APPLICANT: Pathoarray GmbH
; APPLICANT: Stuhlmüller, Bruno
; APPLICANT: Haupt, Thomas
; TITLE OF INVENTION: Nucleic Acid Array
; FILE REFERENCE: 030027US
; CURRENT APPLICATION NUMBER: US/10/278,698
; CURRENT FILING DATE: 2002-10-23
; NUMBER OF SEQ ID NOS: 1050
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 276
; LENGTH: 2947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-278-698-276

Query Match          4.9%; Score 55.6; DB 19; Length 2947;
Best Local Similarity 49.3%; Pred. No. 0.00011;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 325 GCGGTTGGCAGGAGCGCGGAGTCACTGGCTTCTTGGCAGCGCAACCCCTGGAGCTGGA 384
Db 123 GCCCAGCGCGCTGCGCGCGCGCCCATGAGAGCGCGCGCGCTGCGCGCGCGCTCCCGT 182
```

Qy 385 GCCCGCTGCCACCCGGCAGCGGGTGATCTCTGCCCAACCCAGAGAGAGCGCTCTTTC 444
Db 183 TCCGCTCTGCTGCTCGCGCGCTTGGCGCTGCTGCGCGCGAGTGCAGCGGATGTCCT 242
Qy 445 TCGCGCTCGCGCGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTGCGCTGCCCTGCA 504
Db 243 CTTGGAGCCTCTGTCGGACGAGACCGGATGGCCACTCATCAGAGGACTGCTCGCT 302
Qy 505 GCCCTTCTGTACCCAGACACCGCGGATAGGCCCTTTTCAGGCGCAGGCCCCAGAGAGCCT 564
Db 303 GCCATATGCTACGGAATCCAAAGATGAGGATGGTGCAGGAGCAGTGTGCCACGCCA 362
Qy 565 GGACGTGCTGCTCGCGCACCCCTCAGGCTGGTGTGCTGGAGAACGAAACCG 618
Db 363 GCTGGAGGAGCTGCACTGTGCCACGGGCATCAGCCTGGCCAAACGAGCAGGACCG 416

RESULT 14

US-10-278-698-790
; Sequence 790, Application US/10278698
; Publication No. US20050037344A1
; GENERAL INFORMATION:
; APPLICANT: PathoArray GmbH
; APPLICANT: Stuhlmueller, Bruno
; APPLICANT: Haupl, Thomas
; TITLE OF INVENTION: Nucleic Acid Array
; FILE REFERENCE: 030027US
; CURRENT APPLICATION NUMBER: US/10/278,698
; CURRENT FILING DATE: 2002-10-23
; NUMBER OF SEQ ID NOS: 1050
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 790
; LENGTH: 2947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-278-698-790

Query Match 4.9%; Score 55.6; DB 19; Length 2947;
Best Local Similarity 49.3%; Pred. No. 0.00011;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy 325 GCGGTGGACGAGCGCGGAGTACTGCTCTTCGCGACCGAACCCCTGACCTGGA 384
Db 123 GCCACCGCCCGTGCCTCGCGCGCCATGAGCGCGCGCGCTCGCGCGGGTCCCGCT 182
Qy 385 GCGCGCTGCCACCGCGAGCGGGTGATCTCTGCCACCCAGAGAGCAGCTCTTTC 444
Db 183 TCCGCTGCTGCTCGCGCGCTTGGCTGCTGGCGCGGAGTGCACCGGATGTCCT 242
Qy 445 TCGCGTGGCGCGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTGCGCTGCCCTGCA 504
Db 243 CTTGGAGGCTGCTGTGCGGACGACACCGATGGCCACTCATCAGAGGACTGCTCGCT 302
Qy 505 GCCCTTCTGTACCCAGACACCGGGGATAGGCCCTTTTCAGGCGCAGGCCCCAGAGAGCCT 564
Db 303 GCCATATGCTACGGAATCCAAAGATGAGGATGGTGCAGGAGCAGTGTGCCACGCCA 362
Qy 565 GGACGTGCTGCTCGCGCACCCCTCAGGCTGGTGTGCTGGAGAACGAAACCG 618
Db 363 GCTGGAGGAGCTGCACTGTGCCACGGGCATCAGCCTGGCCAAACGAGCAGGACCG 416

RESULT 15

US-10-425-114-31082/c
; Sequence 31082, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31082
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLB73125G09_FLI
US-10-425-114-31082

Query Match 4.8%; Score 54.8; DB 17; Length 1240;
Best Local Similarity 46.2%; Pred. No. 0.00019;
Matches 217; Conservative 0; Mismatches 252; Indels 1; Gaps 1;
Qy 646 CTTGGAGGAGCGCGCCCGGCGCAAGCGCGGAGGAGGCCCTCTCCTAGGAGCAGCGG 705
Db 751 CCAGGACGTAGCAGTCTGCGGCGCGCTTGAGCGCGCGCTGTAGTAGTCTCTCGGCGC 692
Qy 706 TCCCCAGTCTGTGCTTCCCGCGCTACGAGCAGCGCGGAGATGAGCTGTCGCTGCC 765
Db 691 GGCACGCTCGCGCTGGTCTCCCAAACGAGGCCCGCTAGACGACAGCGCTCGGCGT 632
Qy 766 CGCGCGGCGCGCTGCGGCTGTGAAACGTGAGAACGCGCGGCTGTGGCTATGCAGGTA 825
Db 631 CGCGGGGCTGCGGAGGATGGCGGCTCGCAGTAGTACTCTGCGGCCCTTGCAGGCTCCCGG 572
Qy 826 CGCGCAGCGCGCGGCTACTTCCCGCGGTGTGCTGCGCGCGGAGAGGGCTGGGCGCTCT 885
Db 571 CCACCTCTTGAGGAACCTGGCGTAGTGACACGAGGACGAGGTGCGCGGCTCGGCT 512
Qy 886 CTTGAGCGGAGCGGGTTCCGTGGAGAGACGACCGCGGGGTGAGGCGCGGGCTTCCC 945
Db 511 CGATCATCTGCGGTTAGTGGGCGTCCGTGGCGCGGCTGTGTCACCGTTCGCGCTCC 452
Qy 946 TGAACCTCTCCAGGCGCACCGCCCTCCCGCCACCGTGCACCCCGACCTTCGCGCGGCGC 1005
Db 451 CGATCCACCTTTCGCGCGCTGCTGCGCAGCGCTTCCGCGCGCGCGCTGCGCGCGA 392
Qy 1006 CATTCAGAGCCGCTGTGACCGTCAACGCGAGGCGCTTGGAGCGGCGCCACGCGGCCA 1065
Db 391 CGCCCTGTCTCACGAACGCGAGCGCGCGCGGCGGACAGCTCCAGCCCGCTTGTCTCA 332
Qy 1066 G-GGCGCGCTCGAGGCTGCTGGACTCTGTGCGCGCACCCCGACGAG 1114
Db 331 GCAGCGCAGCAGCGCGGACGCTCCACGCGCGCGCGCGCGCGCG 282

Search completed: May 31, 2005, 01:45:34
Job time : 1024 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 18:43:08 ; Search time 320 Seconds
(without alignments)
5788.329 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agcattggaggccccgat.....agcagtgcgagcgatcc 1132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfileel.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	4.4	971	3	US-08-630-915A-197
2	50	4.4	971	4	US-09-879-957-197
3	50	4.4	1272	4	US-09-949-016-3684
C 4	50	4.4	8438	1	US-07-945-283-1
5	50	4.4	42250	4	US-09-949-016-15426
C 6	49.4	4.4	152393	4	US-09-949-016-14514
C 7	49.4	4.4	152393	4	US-09-949-016-14515
C 8	49.4	4.4	156894	4	US-09-949-016-12765
C 9	49.4	4.4	156894	4	US-09-949-016-12766
C 10	49.4	4.4	156895	4	US-09-949-016-16957
C 11	49.4	4.4	156895	4	US-09-949-016-16958
C 12	49.4	4.4	156895	4	US-09-949-016-16959
13	49.2	4.3	801	2	US-08-770-379-16
14	49.2	4.3	801	3	US-08-757-669A-16
15	49.2	4.3	801	3	US-09-230-371A-16
C 16	49.2	4.3	3066	4	US-10-237-551-152
C 17	49.2	4.3	154746	4	US-09-827-688-8
C 18	49.2	4.3	154746	4	US-09-827-688-8
C 19	48.4	4.3	1457	3	US-09-444-053-3
C 20	47.8	4.2	19152	4	US-09-949-016-12110
C 21	47.8	4.2	19153	4	US-09-949-016-15795
C 22	47.8	4.2	90618	4	US-09-949-016-15964
C 23	47.6	4.2	801	3	US-09-298-568-3
C 24	47.6	4.2	801	4	US-09-894-273-3
C 25	47.4	4.2	76985	4	US-09-949-016-12416
C 26	47.4	4.2	76986	4	US-09-949-016-13120
C 27	46.8	4.1	1023	4	US-09-902-540-8340

28	46.8	4.1	2561	4	US-09-616-289-48	Sequence 48, Appl
C 29	46.8	4.1	8578	4	US-09-902-540-871	Sequence 871, App
30	46.6	4.1	1059	4	US-09-902-540-5064	Sequence 5064, Ap
C 31	46.6	4.1	3462	4	US-09-949-016-2642	Sequence 2642, Ap
C 32	46.6	4.1	3641	4	US-09-949-016-3877	Sequence 3877, Ap
C 33	46.6	4.1	7460	4	US-09-949-016-12375	Sequence 12375, A
C 34	46.6	4.1	7462	4	US-09-949-016-14384	Sequence 14384, A
C 35	46.6	4.1	7700	4	US-09-949-016-15619	Sequence 15619, A
C 36	46.6	4.1	28194	4	US-09-902-540-1250	Sequence 1250, Ap
C 37	46.6	4.1	152331	3	US-09-128-155-16	Sequence 16, Appl
C 38	46.4	4.1	1722	4	US-09-902-540-9668	Sequence 9668, Ap
39	46.4	4.1	14555	4	US-09-902-540-1096	Sequence 1096, Ap
40	46.4	4.1	71574	4	US-09-949-016-15580	Sequence 15580, A
41	46	4.1	9369	4	US-10-237-551-190	Sequence 190, App
42	46	4.1	9369	4	US-10-237-551-247	Sequence 247, App
43	45.8	4.0	1221	4	US-09-252-991A-8921	Sequence 8921, Ap
44	45.8	4.0	1545	4	US-09-252-991A-8710	Sequence 8710, Ap
45	45.8	4.0	2199	4	US-09-373-272-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-630-915A-197
; Sequence 197, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mierock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 971 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-08-630-915A-197

Query Match 4.4%; Score 50; DB 3; Length 971;
Best Local Similarity 52.3%; Pred. No. 0.058;
Matches 134; Conservative 0; Mismatches 120; Indels 2; Gaps 1;

597	Qy	GGTGGTGGAGAA	CAGAGACCGCAGAC	CGCGCTGGTTT	CCAGCGCCCT	ACCTCGAGGAGG	656
60	Db	GGCAAGACG	CGCAGGAAGACC	CAGCGCGGGGAT	CGTCCCCAC	CGCCACACGACGCGC	119
657	Qy	CGGCCCCGGG	CCAAAGCCCGGAGG	GAGGCCCGT	CCCTAGGAGCAG	CGGGTCCCCAGTTCT	716
120	Db	GAGTACCCG	CCATGGCAGCGG	CGCCGACCGCAT	CTACG--	ACCTCAACATCCCGGCT	177
717	Qy	GTGCTTCCG	CGCCTACGAGAGC	AGCCGCGCAG	ATGAGTGT	CCGTGCCGCGGGGGCGC	776
178	Db	TCGTCAA	GTTCGCTAT	GTGGCCGAGCGG	GAGGATGAGTTGT	CCCTGGTGAAGGGTCCG	237
777	Qy	GGGTGGCGGT	GTGTGAAACGTC	CAGACCGCGCT	CGTGGCTAT	GCAGGTACGGGACCGGG	836
238	Db	GGTCA	CGTCA	TGAGAG	AGTGCAGCG	GTTCGTGGCGGGCAGCT	297
837	Qy	CGGGCCT	ACTCCCCG	852			
298	Db	TCGGCTG	GTTCCTC	313			

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

Query Match 4.4%; Score 50; DB 4; Length 971;
 Best Local Similarity 52.3%; Ref. No. 0.058;
 Matches 134; Conservative 0; Mismatches 120; Indels 2; Gaps 1;

QY	597	GGCTGGTGGAGAAAGACGAGACGGGAGACCGCGCTCGTGTTCACAGCGCCCTACCTGGAGGAGG	656
Db	60	GGCAAGACGCGCAGAGACAGCGCGCGGATGGTCCCCACGCGCCACGACGACGCGC	119
QY	657	CGGCCCCGGGGCCAAAGGCGGGAGGAGGCGCCCGTCCCTAGGAGGAGCAGCGGTCCCCAGTTCT	716
Db	120	GAGTACCCCGCATGGCAGCGCGCGCGACCGCATCTACG--ACCTCAACATCCCGGCT	177
QY	717	GTGCTTCCCGGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCGCGGGGGCGC	776
Db	178	TCGTCAAGTTCGCGCTATGTGCCGAGCGGAGGATGAGTTCTCCCTGGTGAAGGGGTGCG	237
QY	777	CGGTGCGGCTTTCGAAACGTACAGACCGCGCGCTGGTGGCTATGCGGTACGGCGACCGGG	836
Db	238	CGGTACCGCTCATGAGAAGTGCAGCGACGCGTGGTGGCGGGCAGCTACAACGGGCAGA	297
QY	837	CGGGCGCTACTCCCGCGC	852
Db	298	TCGGTGGTCCCCCTC	313

RESULT 3
 US-09-949-016-3684
 ; Sequence 3684, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3684
 ; LENGTH: 1272
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-3684

Query Match	4.4%;	Score 50;	DB 4;	Length 1272;
Best Local Similarity	52.3%;	Pred. No. 0.061;		
Matches 134;	Conservative 0;	Mismatches 120;	Indels 2;	Gaps 1;
Qy	597	GGCTGTGGAGAACGAAGACCGGCAGACCGCCTGTTTCCAGCGCCTACTCTGGAGGAGG	656	
Db	256	GGCAAGACGCGCAGGAAGACAGCGCGCGGATCGTCCCGCACGCCAGCACGACGCC	315	
Qy	657	CGGCCCGCGCCRAAGCGCGGAGGAGCGCGTCCCTAGGAGAGCAGCGTCCCGCAGTTCT	716	
Db	316	GAGTACCCCGCCAAATGGACGCGCGCCAGCCGCACTACG--ACCTCAACATCCCGGCT	373	
Qy	717	GTGCTTCCCGCCTACGAGAGACGCGCAGATGAGCTGTCGTGCCCGGGGGCGC	776	
Db	374	TCGTCAAGTTGCGCTATGTGCCAGCGGAGGATGAGTTGCTCCTGTTGAAGGGGTGCG	433	
Qy	777	CGGTGCGCGTGTGGAAACGTCAGACCGCGCTGGTGGCTATGCAGGTACCGGACCGGG	836	
Db	434	CGGTACCGTCATGAGAGGTGCAGGCAGCGTTCGTTGGCGGGCGAGCTACAACGGCAGA	493	
Qy	837	CGGGCCTACTCCCGC	852	
Db	494	TCGGCTGTTCCCTC	509	

```

RESULT 4
US-07-945-283-1/c
; Sequence 1, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LLT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8438 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 622..6495
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1099, "g")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1267, "t")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1381, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(1566, "c")
; FEATURE:
; NAME/KEY: variation
; LOCATION: replace(7010, "g")
;
US-07-945-283-1

Query Match 4.4%; Score 50; DB 1; Length 8438;
Best Local Similarity 44.1%; Pred. No. 0.09;
Matches 299; Conservative 0; Mismatches 375; Indels 3; Gaps 2;

QY 452 GCGGGCGGCTCTCCATCCAGTCTGGAGCTCAGAGCTGGCGCTGCAGCCCTTC 511
Db 4695 GCTTGGCGGGCGTCTTCGAGCGCTGCTGCTGCGCGCTGCGCGCGCGCC 4636

```

Qy	597	GGCTGTGGAGAACGAGACCGGCAGACCGCCTGTGTTTCCAGCGCCTTACCTGGAGGAGG	656
Db	28310	GGCAAGACGCGCAGGAAGACCAAGCGCGGGATGCGTCCCCACGCCAGCACGACGC	28369
Qy	657	CGGCCCCGGGCCAAGCGCGGGAGGGAGGCCGCTCCCTAGGGAGCAGCGGTCCCCAGTTCT	716
Db	28370	GAGTACCCCGCCCAATGGCAGCGCGCGACCGCATCTACG--ACCTCAACATCCCGGCT	28427
Qy	717	GTGCTTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCGTTGCCCGGGGGCGC	776
Db	28428	TCGTCAAGTTCGGCTATGTGGCCAGCGGAGATGAGTTGTCCCTGTGTAAGGGGTGCG	28487
Qy	777	CGCTGGCGGTGTTTGGAAACGTCAAGACCGCGGCTGTGTGGCTATGCAAGTACGCGCACCGGG	836
Db	28488	CGCTACCGTGTGAAGAGAGTGCAGCGACGGTTGGTGGCGGGCAGCTACACCGGCAGA	28547
Qy	837	CGGGCCTACTTCCCGC	852
Db	28548	TCGCGTGGTTCCCGCTC	28563

```

RESULT 6
US-09-949-016-14514/c
; Sequence 14514, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14514
; LENGTH: 152393
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(152393)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14514

```

Db	2280	CCTGCCCGCGCCTTATCAATAACGCCGCCA - CGCGGGGTTCGAGACCGCGCGCCCGCCGCGCA	2222
Qy	769	GGGGCGCGCGCTGCGCGCTGTGTGGAACGTCAGACCGCGCGCTGCTGGCTATGCAGGTACGG	828
Db	2221	GC CGCGCGCGCGCGCGCGCGCTCACTTACGCGCAGCGCGCGCGCTGGTCTCTCG	2162
Qy	829	CGACCGCGCGCGCGCTACTCCCGCGGTGCTGTCGGCGCGGAGAGGGCTGGCGCTCTCT	888
Db	2161	CACGCGCGCGCGCGCGCTTCCGCGCGAGTCTGCGCGCGCGCGCGCTCTGCTCTGTTGTT	2102
Qy	889	GAGCGGAGCGGGTTCCTGAGAGAGACCCGCGCGGTGAGGCCCGGG - GCTTCCCT	946
Db	2101	GTGCGAGCGCGCGCGGTGCGCGTGGCTTCTGCGCGCGCGCGAGCGCGCGCGCACT	2042
Qy	947	GAACCTCTCCAGGCGACCGCGCTCCCCCGACCG - TGCCACCCGACCTTCGCGGGGCG	1005
Db	2041	GGCGCTTGGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAGCGCT	1982
Qy	1006	CATCAGAGCGCTGTGTCACCGTCACGAGGGCCCTGAGAGGGCGCGCGCGCGCGCA	1065
Db	1981	CAGCGCGCGCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1922
Qy	1066	GGCGCGCGCTCGAGGTGCGTGGACTCTGTCCCGCAC	1102
Db	1921	ACGCGAGGGGCGGGGCGCGGACGCGCGGAC	1885

```

RESULT 7
US-09-949-016-14515/c
; Sequence 14515, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ IDS NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14515
; LENGTH: 152393
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(152393)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14515

```

✓

Db 2340 CTTCTCCGCGGCGCCCTCCGCTCCCGGAGCGGCCGACAGCTCCGGGCCCCACCAACAGTC 2281
Qy 709 CCAGTTCTGTCTTCCCGCCCTACAGAGCAGCCGCGCAGATGAGCTGTCCGTGCCCCG 768
Db 2280 CTTGCCCGCGCTTACATAAACGCCGCCA-CGCGGGGTCCGAGCCCGCGCCGCCGA 2222
Qy 769 GGGGGCGCGGTGCGGTGTGGAACGTGAGACCGCGGTGTGGCTATGCAAGGTACGG 828
Db 2221 GCGCGCGCGCCCGCGCCCTCCTACCTCAGCGCAGCCGCGCGCGCTGTCTCTCCG 2162
Qy 829 CGACCGGGGGGCTACTCTCCCGGTGTCTGCGGCCGGAAGGGTGGCGCTCTCCT 888
Db 2161 CAGCGCGCCCAAGCTTCGCGCGAGCTCTGCGCCGGGCGGCTCTGCTCTGTGGTT 2102
Qy 889 GAGCGGAGCGGGTTCCGTGGAGGAGACGACCCGCGGGGTGAGGCGCCGGG--GCTTCCCT 946
Db 2101 GTGCGCAGCGCGCGGTGCGGTGCGCTGCGCGGCCCGAGCCCGCGCCCACT 2042
Qy 947 GAACCTTCCAGCGCACCGCCCTTCCCGCCACCG-TGCCACCGACCTTTCGCGGGCGC 1005
Db 2041 GGCCTGTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCCACCGCGCTCAGCGCT 1982
Qy 1006 CATCCAGAGCGCTGTGCACTGTACACGACGCGGCTTGAGCGGGGCCCAAGGCCCA 1065
Db 1981 CAGCCCGCGCGCGCGGTCTCCCGCCCTCTCCCGCCCGCGCGCGCGCGCGCGC 1922
Qy 1066 GGGCGCGCTCGAGGGTGTGACTCTGTGCGCGCAC 1102
Db 1921 ACGCCAGGGGGCGGGCGCGCAGCGACCGCGAC 1885

RESULT 8
US-09-949-016-12765/c
; Sequence 12765, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12765
; LENGTH: 156894
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(156894)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-12765

Query Match 4.4%; Score 49.4; DB 4; Length 156894;
Best Local Similarity 46.5%; Pred. No. 0.22;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
Qy 470 CACAGTCTGAGGCTCAGAGCTTGCCTGCTGAGCCCTTGTATCCAGGACACGCG 529
Db 2520 CGCAGCTCTGCTCCCGCGCAGCTTCCCTCCCGCGCGAGCGCGCATCCACGCG 2461
Qy 530 GATAGGCTTTTCAGGCGGAGGAGAGGCTGAGC-TGCTCTGCGGACCCCTC 588
Db 2460 ACTCGGCGCCCTTCGAGAGGGAGCGCGGGCGCGCGCGAGCTTCTCCCGGAGCCGA 2401

Qy 589 AGCTGTGCTGTGAGAAAGAACCCGCGCAGACCGCTGTGTTTCCAGCGCCCTACCT 648
Db 2400 GCSCCGAGCAGACGGCGGACCCAGAGGGTGGGTGCGAGCGGTGCGCGGACCTCG 2341
Qy 649 GAGGAGGGGGCCCCCGGCGCAAGCGCGGAGGAGGAGCCCTTCCCTAGGAGAGCAGCGTCC 708
Db 2340 CTTTCTCCGGGGGGCGCTTCCGCTCCCGGAGACGGCCCGCAGCTCCGGCCCCACCAACAGTC 2281
Qy 709 CCAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCCGCGCAGATGAGCTGTCCGTGCCCCG 768
Db 2280 CTTGCCCGCGCTTACATAAACGCCGCCA-CGCGGGGTCCGAGCCCGCGCCGCCGA 2222
Qy 769 GGGGGCGCGGTGCGGTGTGGAACGTGAGACCGCGGTGTGGCTATGCAAGGTACGG 828
Db 2221 GCGCGCGCGCCCGCGCGCTCCTACCTCAGCGCAGCGCGCGCGCTGTGCTCTCCG 2162
Qy 829 CGACCGGGGGGCTACTCTCCCGGTGTCTGCGGCCGGAAGGGTGGCGCTCTCCT 888
Db 2161 CAGCGCGCCCAAGCTTCGCGCGAGCTCTGCGCCGGGCGGCTCTGCTCTGTGGTT 2102
Qy 889 GAGCGGAGCGGGTTCCGTGGAGGAGACGACCCGCGGGGTGAGGCGCCGGG--GCTTCCCT 946
Db 2101 GTGCGCAGCGCGCGGTGCGGTGCGCTGCGCGGCCCGAGCCCGCGCCCACT 2042
Qy 947 GAACCTTCCAGCGCACCGCCCTTCCCGCCACCG-TGCCACCGACCTTTCGCGGGCGC 1005
Db 2041 GGCCTGTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCCACCGCGCTCAGCGCT 1982
Qy 1006 CATCCAGAGCGCTGTGCACTGTACACGACGCGGCTTGAGCGGGGCCCAAGGCCCA 1065
Db 1981 CAGCCCGCGCGCGGTCTCCCGCCCTCTCCCGCCCGCGCGCGCGCGCGCGC 1922
Qy 1066 GGGCGCGCTCGAGGGTGTGACTCTGTGCGCGCAC 1102
Db 1921 ACGCCAGGGGGCGGGCGCGCAGCGACCGCGAC 1885

RESULT 9
US-09-949-016-12766/c
; Sequence 12766, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 207012
; SEQ ID NO 12766
; LENGTH: 156894
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(156894)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-12766

Query Match 4.4%; Score 49.4; DB 4; Length 156894;
Best Local Similarity 46.5%; Pred. No. 0.22;
Matches 296; Conservative 0; Mismatches 336; Indels 5; Gaps 4;
Qy 470 CACAGTCTGAGGCTCAGAGCTTGCCTGCTGAGCCCTTGTATCCAGGACACGCG 529
Db 2520 CGCAGCTCTGCTCCCGCGCAGCTTCCCTCCCGCGCGAGCGCGCATCCACGCG 2461


```
QY 534 GGCTTTTTCAGGCGCAGGCCCCAGAGAGCCTGACGCTGCTGTCGCGACACCCCTCAGGCT 593
DB 213 CGCCAGCGCGCGCAGCGCCCGCGGCGCCCGAGCCCGAGCGGCGCGGCTACGGGCT 272
QY 594 GGTGGCTGTGGAGACGAAGACCGGACAGCCGCTGCTGTTTCAGCGCCCTACTCTGAGG 653
DB 273 AGGCCACGAATAATTTTTTTTCGGGCGGCCCCCGAACCTCTCTCGGCCCCCCCGGTCCC 332
QY 654 AGCGGCCCCCGGC-CAAGGCGCGGAGGAGGCGCCGTCCTTAGGAGCAGCGGTCCCCAG 712
DB 333 CGCGGCGCGCGCGCCCCCGGGGGGTAAACAGAGGGGGGGGGATGCGGCGCGGC 392
QY 713 TTCTGTGCTTCCGCGCTACAGAGAGCAGCGCGCAGATGAGTGTCTCGTGCCTGAGGG 772
DB 393 GCGCGCGCGCGCGCGCGCTTGTCTTCTGTTTCTCCGCGGCCCCCGCGCGCAG 452
QY 773 GCGCGGTGCGCGTGTGGAACGTACAGACCGCGGCTGTGGCTATGACAGGTACGGGAC 832
DB 453 CGCGCGCGCGCGCGCGCCCCCTCCCGCGGGGCTCGCGGGGGGCCCCCTGTCCC 512
QY 833 CGCGCGCGCTACTCCCGCGGTGCTGTCGCGCGCGAAGGGCTGGCGCTCTCTGAGC 892
DB 513 CGCGCGCGCGCGACCCCCCGCGCGCGCGCGCGCGATCCCGCGGCGCGCGCCCC 572
QY 893 GGGACCGGGTTCCGTGGAGAGACGCCCGCGGGTGAAGCCCGGGCTTCCCTGAACCC 952
DB 573 TGCGGGGACGCGCGCGCGCTCGCGCGCTCCCGCGGATGGGCGCGCGCGCGCC 632
QY 953 TCCAGGCACCGCCCCCTCCCGACGTCGCCACCGACCTTCGCGCGGCGCGCATCCAG 1012
DB 633 TCAGGCGCGCGCGCGCGCGCTGTCTCCCGCCCCCGCGCGGGAACCCCGGCGAG 692
QY 1013 AG 1014
DB 693 CG 694

RESULT 15
US-09-230-371A-16
; Sequence 16, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-16
```

```
Query Match 4.3%; Score 49.2; DB 3; Length 801;
Best Local Similarity 44.1%; Pred.No.0.084;
Matches 292; Conservative 0; Mismatches 368; Indels 2; Gaps 2;

QY 354 GCTTCTTCGACCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCCGGCGAGCGGGTGA 413
DB 34 GCGCGCGCCCTCCCTCCCTCCCGCTCGCTCCCGCGCTGCGCGCGAGCCCGCGCGGA 93
QY 414 TCGTGGCCACCCAGAGAGACGCTCTTTCTGCGCGCTCGGGCGCGCTCTCCATCACA 473
```

```
DB 94 GCGGCGCGCCCGCGGGGGGAGGGCGCGCCCGCGGCTCCCTGCGGGGGCGGGGAGCGG 153
QY 474 GTCTGAGGCTCAGAGCCTGCGCTGCGAGCCCTTCTGTATCCCAAGGACACGCGGGATA 533
DB 154 GAGGGGGGCGCGGCGCCCGCGCGCGCGGAG-CGCGAGGGGCCCCCGCGCGC 212
QY 534 GGCCTTTTTCAGGCGCAGGCCCCAGAGAGCCTGAGACGTGCTGTCGCGGACCCCTCAGGCT 593
DB 213 CGCCAGCGCGCGCAGGCCCCCGGGGCGCGAGCCCGCGGCGCGGGGTACGGGGCT 272
QY 594 GGTGGCTGTGGAGAACGAAGACCGGACAGCCGCTGTTTTCAGCGCCCTACTCTGAGG 653
DB 273 AGGCCACGAATAATTTTTTTTCGGGCGGCCCCCGAACCTCTCTCGGCCCCCCCGGTCCC 332
QY 654 AGCGGCCCCCGGC-CAAGGCGCGGAGGAGGCGCCGTCCTTAGGAGCAGCGGTCCCCAG 712
DB 333 CGCGGCGCGCGCGCCCCCGGGGGGTAAACAGAGGGGGGGGGATGCGGCGCGGC 392
QY 713 TTCTGTGCTTCCGCGCTACAGAGAGCAGCGCGCAGATGAGTGTCTCGTGCCTGAGGG 772
DB 393 GCGCGCGCGCGCGCGCGCTTGTCTTCTGTTTCTCCCGCGGCCCCCGCGCGCAG 452
QY 773 GCGCGGTGCGCGTGTGGAACGTACAGACCGCGGCTGTGGCTATGACAGGTACGGGAC 832
DB 453 CGCGCGCGCGCGCGCGCGCCCCCTCCCGGGGCTCGCGGGGGGCCCCCTGTCCC 512
QY 833 CGCGGCGGCTACTCCCGCGGTGCTGTCGCGCGCGAAGGGCTGGCGCTCTCTGAGC 892
DB 513 CGCGCGGCGCGCGACCCCCCGCGCGCGCGCGCGATCCCGCGGCGCGCGCCCC 572
QY 893 GGGACCGGGTTCCGTGGAGAGACGACCGCGGGTGAAGCCCGGGCTTCCCTGAACCC 952
DB 573 TGCGGGGACGCGCGCGCGCTCGCGCGCTCCCGCGGATGGGCGCGCGCGCGCC 632
QY 953 TCCAGGCACCGCCCCCTCCCGACCGTCGCCACCGACCTTCGCGCGGCGCGCATCCAG 1012
DB 633 TCAGGCGCGCGCGCGCGCGCTGTCTCCCGCCCCCGCGCGGGAACCCCGGCGAG 692
QY 1013 AG 1014
DB 693 CG 694
```

Search completed: May 30, 2005, 23:21:52
Job time : 324 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 15:46:02 ; Search time 973 Seconds
(without alignments)
6887.099 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agcattggcagggcccccgcatt.....agcattggcagggcagggatcc 1132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn19908:*
- 2: Geneseqn19908:*
- 3: Geneseqn20008:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1132	100.0	1132	12	ADJ57070 Coding se
2	1116	98.6	1129	12	ADJ57068 Coding se
3	1107	97.8	1147	12	ADJ57072 Coding se
4	1101.8	97.3	1520	13	ADQ85428 Human tum
5	1091	96.4	1144	12	ADJ57074 Coding se
6	798	70.5	1911	5	AAS89078 Human encod
7	798	70.5	2881	5	AAS89434 Human encod
8	342.4	30.2	2922	11	ADM02780 Human cdn
9	77.2	6.8	114955	2	AAX53491 Human ade
10	71.4	6.3	114955	2	AAX53491 Human ade
11	56.8	5.0	2781	5	ABV29190 Human pro
12	56.8	5.0	2781	5	ABV29190 Human pro
13	55.6	4.9	1850	13	ABD33012 Human can
14	55.6	4.9	2196	13	ACN42908 Human dia
15	55.6	4.9	2215	13	ABD33013 Human can
16	55.6	4.9	2255	13	ABD33018 Human can
17	55.6	4.9	2422	13	ACN42907 Human dia
18	55.6	4.9	2445	13	ABD33016 Human can
19	55.6	4.9	2448	13	ACN42906 Human dia
20	55.6	4.9	2453	13	ACN42913 Human dia

21	55.6	4.9	2470	13	ACN42912	Human dia
22	55.6	4.9	2579	13	ACN42911	Human dia
23	55.6	4.9	2580	9	AAL62035	Human cel
24	55.6	4.9	2622	13	ABD33019	Human can
25	55.6	4.9	2847	13	ACN42905	Human dia
26	55.6	4.9	2909	13	ABD33017	Human can
27	55.6	4.9	3004	13	ACN42910	Human dia
28	55.6	4.9	3121	13	ACN42909	Human dia
29	54.8	4.8	4667	12	ADQ22939	Human sof
30	54.8	4.8	84428	12	ADM45913	Streptomy
31	54	4.8	1574	3	AAZ48807	Rice inob
32	53.2	4.7	4545	10	ADE52577	Human SQV
33	53.2	4.7	4545	13	ABD33445	Human can
34	53.2	4.7	95914	13	ABD33444	Human can
35	52.8	4.7	2529	12	ADH18908	Human cel
36	52.8	4.7	2763	12	ADH18909	Human cel
37	52.8	4.7	110000	12	ADQ97138_3	Conitination (4 of
38	51.8	4.6	2156	5	AAS83572	Human fib
39	51.8	4.6	2156	6	AAZ4870	Human fib
40	51.8	4.6	2156	12	ADN03778	Human fib
41	51.8	4.6	2200	2	AAQ11010	Fibulin C
42	51.8	4.6	2349	6	AAZ4868	Human fib
43	51.8	4.6	2355	2	AAQ11008	Fibulin A
44	51.8	4.6	2359	5	AAS83573	DNA encod
45	51.8	4.6	2359	6	AAZ4871	Human fib

ALIGNMENTS

RESULT 1
ADJ57070
ID ADJ57070 standard; cdna; 1132 BP.
XX
AC ADJ57070;
XX
DT 06-MAY-2004 (first entry)
XX
DE Coding sequence for p41Nox, a regulatory protein for Nox enzyme.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytosstatic; chromosome 16;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 5..1120
FT /*tag= a
FT /product= "Human p41Nox"
XX
PN WO2004007689-A2.
XX
PD 22-JAN-2004.
XX
PF 16-JUL-2003; 2003WO-US022246.
XX
PR 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
XX (UYEM-) UNIV EMORY.
XX
XX Lambeth JD, Cheng G;
DR WPI; 2004-122927/12.
DR P-PSDB; ADJ57071.
XX
XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.
XX
PS Claim 8; SEQ ID NO 3; 70pp; English.
XX

CC The present sequence is that of cDNA encoding a regulatory protein.
CC denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of
CC p41Nox was used as query in database screenings, and a genomic sequence
CC was identified from human chromosome 16. PCR primers based on this
CC sequence were used in RT-PCR experiments to look for expression in a
CC series of normal and cancer cell lines. A 350 bp PCR product was detected
CC in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.
CC RNA from T84 cells and human foetal liver was then used to amplify the
CC full-length p41Nox cDNA by PCR. 13 unique clones were obtained.
CC Sequencing revealed 4 p41Nox splice variants, of which nucleotide and
CC protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins
CC function in the regulation of cell growth and are therefore implicated in
CC diseases involving abnormal cell growth, such as cancer. They may also
CC function in innate immune mechanisms of epithelial tissue or other
CC barrier cells, and hence may be involved in diseases of diminished
CC ability to fight infections or inflammatory conditions. The p41Nox
CC regulatory proteins and nucleic acids are useful for preparing a
CC medicament for affecting superoxide formation or regulating cellular
CC proliferation in an animal or a human. The regulatory proteins are also
CC useful in drug development, e.g. screening for drugs that regulate the
CC biological activity of the proteins, or in assays that relate to
CC assessment of abnormal growth or cellular proliferation including cancer.
XX
SQ Sequence 1132 BP; 173 A; 399 C; 391 G; 169 T; 0 U; 0 Other;

Query Match 100.0%; Score 1132; DB 12; Length 1132;
Best Local Similarity 100.0%; Pred. No. 2.5e-203; Indels 0; Gaps 0;
Matches 1132; Conservative 0; Mismatches 0;

QY 1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTCAAGGGGAGCCCTGGTGACATCAA 60
DB 1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTCAAGGGGAGCCCTGGTGACATCAA 60
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACTTCGTGCG 120
DB 61 GAGGCTCCAAACGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACTTCGTGCG 120
QY 121 CAGGATTTGGGACGAATTCAGGCAGCTCAGAGAGCCCTCAAGGAGACTTCCCGGTGA 180
DB 121 CAGGATTTGGGACGAATTCAGGCAGCTCAGAGAGCCCTCAAGGAGACTTCCCGGTGA 180
QY 181 GCGGGCTGTCTGCGAGATCTGACGGGTTCTCCAAAGCTTCTCGATGACCACTGT 240
DB 181 GCGGGCTGTCTGCGAGATCTGACGGGTTCTCCAAAGCTTCTCGATGACCACTGT 240
QY 241 GGGACGCTGGGGCGCAGAGCCGGCTGCGCGCTGCGAGCTGTGTTGGAAACCTATTTC 300
DB 241 GGGACGCTGGGGCGCAGAGCCGGCTGCGCGCTGCGAGCTGTGTTGGAAACCTATTTC 300
QY 301 TCGGAGGCTGTGGCGACTGACAGCGGTGCGACGGAGCCGACGATCATGGCTTCT 360
DB 301 TCGGAGGCTGTGGCGACTGACAGCGGTGCGACGGAGCCGACGATCATGGCTTCT 360
QY 361 CGCAGCGAACCCCTGGACCTGGAGCCCGCTGCGCTGACCCCGGCGCGGTGATCTGCC 420
DB 361 CGCAGCGAACCCCTGGACCTGGAGCCCGCTGCGCTGACCCCGGCGCGGTGATCTGCC 420
QY 421 CACCCACGAGGAGCAGCTCTTTTCTGCGCTGCGGGCCGCTTCTCCATCCACAGTCTGA 480
DB 421 CACCCACGAGGAGCAGCTCTTTTCTGCGCTGCGGGCCGCTTCTCCATCCACAGTCTGA 480
QY 481 GGCTCAGAGCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTT 540
DB 481 GGCTCAGAGCCCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTT 540
QY 541 TCAGGCGCAGGCCCGCAGGAGCCTGAGCTGTGCTGCGGGCACCCCTCAGGCTGGTGGCT 600
DB 541 TCAGGCGCAGGCCCGCAGGAGCCTGAGCTGTGCTGCGGGCACCCCTCAGGCTGGTGGCT 600
QY 601 GGTGGAGAACGAAGACCGGACAGCCGCTGTGTTTCCAGCGCCCTACCTGGAGAGGCGGC 660
DB 601 GGTGGAGAACGAAGACCGGACAGCCGCTGTGTTTCCAGCGCCCTACCTGGAGAGGCGGC 660

QY 661 CCCGGGCCCAAGCGCGGAGGAGCGCCCTCTCTAGGAGAGCAGCGTCCCGCTTCTGTGC 720
DB 661 CCCGGGCCCAAGCGCGGAGGAGCGCCCTCTCTAGGAGAGCAGCGTCCCGCTTCTGTGC 720
QY 721 TTCCCGCGCTTACGAGAGCAGCCCGCGCAGATGAGCTGTCCCTGCGGGGCGCGCT 780
DB 721 TTCCCGCGCTTACGAGAGCAGCCCGCGCAGATGAGCTGTCCCTGCGGGGCGCGCT 780
QY 781 GCGGTGTGTAAGAGCTCAGACCGGGCTGGTGGCTATGACAGTACGGGACCGGGCGG 840
DB 781 GCGGTGTGTAAGAGCTCAGACCGGGCTGGTGGCTATGACAGTACGGGACCGGGCGG 840
QY 841 CCTACTCCCGCGTGTGCTGCGCGCGAAGGGCTGGCGCTCTCTCAGAGCGGACGG 900
DB 841 CCTACTCCCGCGTGTGCTGCGCGCGAAGGGCTGGCGCTCTCTCAGAGCGGACGG 900
QY 901 GTTCCGTGGAGAGCAGACCCCGCGGGGTGAGCGCGGGCTTCCCTGAACCTCCAGGC 960
DB 901 GTTCCGTGGAGAGCAGACCCCGCGGGGTGAGCGCGGGCTTCCCTGAACCTCCAGGC 960
QY 961 CACGCGCCCTCCCGCAGCGTGGCCACCCGACCTTCCGCGGGCGGCTCAGAGCGCTG 1020
DB 961 CACGCGCCCTCCCGCAGCGTGGCCACCCGACCTTCCGCGGGCGGCTCAGAGCGCTG 1020
QY 1021 CTGCACGCTCACAGCGGGCCCTGGAGCGCGCCACCGCGCGCGGCGCCCTCGAGG 1080
DB 1021 CTGCACGCTCACAGCGGGCCCTGGAGCGCGCCACCGCGCGCGGCGCCCTCGAGG 1080
QY 1081 GTGCGTGGACTCTGTGCGCGCACCCCGACGAGCAGTGAGCGGAGGATCC 1132
DB 1081 GTGCGTGGACTCTGTGCGCGCACCCCGACGAGCAGTGAGCGGAGGATCC 1132

RESULT 2

ADJ57068
ID ADJ57068 standard; cDNA; 1129 BP.
XX
AC ADJ57068;
XX
DT 06-MAY-2004 (first entry)
XX
DE Coding sequence for p41Nox, a regulatory protein for Nox enzyme.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; chromosome 16;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 5..1117
FT /*tag= a
FT /product= "p41Nox"
XX
PN WO2004007689-A2.
XX
PD 22-JAN-2004.
XX
PF 16-JUL-2003; 2003MO-US022246.
XX
PR 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Lambeth JD, Cheng G;
XX
DR WPI; 2004-122927/12.
DR P-PSDB; ADJ57069.
XX
XX
PT New regulatory protein for hydrogenerated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.

XX	Claim 8; SEQ ID NO 1; 70pp; English.	
PS		
XX		
CC	The present sequence is that of cDNA encoding a regulatory protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of p47phox was used as query in database screenings, and a genomic sequence was identified from human chromosome 16. PCR primers based on this sequence were used in RT-PCR experiments to look for expression in a series of normal and cancer cell lines. A 350 bp PCR product was detected in RNA from human foetal kidney, liver, lung and T84 colon cancer cells. RNA from T84 cells and human foetal liver was then used to amplify the full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.	
CC	Sequencing revealed 4 p41Nox splice variants, of which nucleotide and protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins function in the regulation of cell growth and are therefore implicated in diseases involving abnormal cell growth, such as cancer. They may also function in innate immune mechanisms of epithelial tissue or other barrier cells, and hence may be involved in diseases of diminished ability to fight infections or inflammatory conditions. The p41Nox regulatory proteins and nucleic acids are useful for preparing a medicament for affecting superoxide formation or regulating cellular proliferation in an animal or a human. The regulatory proteins are also useful in drug development, e.g. screening for drugs that regulate the biological activity of the proteins, or in assays that relate to assessment of abnormal growth or cellular proliferation including cancer.	
XX		
XX	Sequence 1129 BP; 171 A; 399 C; 390 G; 169 T; 0 U; 0 Other;	
XX		
CC	Query Match 98.6%; Score 1116; DB 12; Length 1129;	
CC	Best Local Similarity 99.7%; Pred. No. 3.5e-210;	
CC	Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;	
QY	1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTGAAGGGGAGCCCTGGTGCAGATCAA 60	
DB	1 AGCCATGGCAGGCCCCCGATACCCAGTTTTCAGTGAAGGGGAGCCCTGGTGCAGATCAA 60	
QY	61 GAGGCTCCTAAAGTTTGGCTTCTCTGTGGCTGGTTCAGACGCGCAGCACCTTCGTGGC 120	
DB	61 GAGGCTCCTAAAGTTTGGCTTCTCTGTGGCTGGTTCAGACGCGCAGCACCTTCGTGGC 120	
QY	121 CAGGAGCTGGGACGAATTCAGGACGCTCAAGAAGACCCTCAAGGAGACCTTCCCGGTGGA 180	
DB	121 CAGGAGTTGGGACGAATTCAGGACGCTC---AAGACCCTCAAGGAGACCTTCCCGGTGGA 177	
QY	181 GGCGGCTGTGTCGGAGATCTGACCGGCTTCTCCCAAAGCTTCTCGATGCACCACTGTT 240	
DB	178 GGCGGCTGTGTCGGAGATCTGACCGGCTTCTCCCAAAGCTTCTCGATGCACCACTGTT 237	
QY	241 GGGAACGCTGGGGCGGACGAGCCGGGCTGGCGGCTTCGAGCTGTTGGAAACCTATTC 300	
DB	238 GGGAACGCTGGGGCGGACGAGCCGGGCTGGCGGCTTCGAGCTGTTGGAAACCTATTC 297	
QY	301 TCGGAGCTGTGGCGACCTGACGAGCGGTGGCAGGAGCCCGGATCATCTGGCTTCTT 360	
DB	298 TCGGAGCTGTGGCGACCTGACGAGCGGTGGCAGGAGCCCGGATCATCTGGCTTCTT 357	
QY	361 CGCACCGCAACCCCTGGACCTCGAGCCCGGCTGTCACCCCGGAGCCGGGTGATCTTGCC 420	
DB	358 CGCACCGCAACCCCTGGACCTCGAGCCCGGCTGTCACCCCGGAGCCGGGTGATCTTGCC 417	
QY	421 CACCCAGAGAGACGACCTCTTTTCTGGCTGCGGGCCGCTCTCCATCCACAGCTTGA 480	
DB	418 CACCCAGAGAGACGACCTCTTTTCTGGCTGCGGGCCGCTCTCCATCCACAGCTTGA 477	
QY	481 GGCTCAGAGCTGCGCTGCTGACGCCCTTCTGTACCCAGGACACGGGGATAGCCCTTT 540	
DB	478 GGCTCAGAGCTGCGCTGCTGACGCCCTTCTGTACCCAGGACACGGGGATAGCCCTTT 537	
QY	541 TCAGGCGCAGGCGCCAGGAGAGCTTGACGTGCTGCTGGCGCACCCCTCAGGCTGGT 600	
DB	538 TCAGGCGCAGGCGCCAGGAGAGCTTGACGTGCTGCTGGCGCACCCCTCAGGCTGGT 597	
QY	601 GGTGGAGAACGAAGACCGGACAGCCGCTGGTTTCCAGCGCCCTTACCTGGAGGAGCGGC 660	

PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
XX a human.

PS Claim 9; SEQ ID NO 5; 70pp; English.

XX
CC The present sequence is that of cDNA encoding a regulatory protein.
CC denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of
CC p47phox was used as query in database screenings, and a genomic sequence
CC was identified from human chromosome 16. PCR primers based on this
CC sequence were used in RT-PCR experiments to look for expression in a
CC series of normal and cancer cell lines. A 350 bp PCR product was detected
CC in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.
CC RNA from T84 cells and human foetal liver was then used to amplify the
CC full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.
CC Sequencing revealed 4 p41Nox splice variants, of which nucleotide and
CC protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins
CC function in the regulation of cell growth and are therefore implicated in
CC diseases involving abnormal cell growth, such as cancer. They may also
CC function in innate immune mechanisms of epithelial tissue or other
CC barrier cells, and hence may be involved in diseases of diminished
CC ability to fight infections or inflammatory conditions. The p41Nox
CC regulatory proteins and nucleic acids are useful for preparing a
CC medicament for affecting superoxide formation or regulating cellular
CC proliferation in an animal or a human. The regulatory proteins are also
CC useful in drug development, e.g. screening for drugs that regulate the
CC biological activity of the proteins, or in assays that relate to
CC assessment of abnormal growth or cellular proliferation including cancer.

XX Sequence 1147 BP; 175 A; 404 C; 397 G; 171 T; 0 U; 0 Other;

Query Match 97.8%; Score 1107; DB 12; Length 1147;
Best Local Similarity 98.7%; Pred. No. 2.1e-208;
Matches 1132; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

QY 1 AGCCATGCGAGCCGCCGATACCCAGTTTCAAGTCAAGGGGAGCCCTGGTGCGATCAA 60
DB 1 AGCCATGCGAGCCGCCGATACCCAGTTTCAAGTCAAGGGGAGCCCTGGTGCGATCAA 60
QY 61 GAGGCTCAAAAGTTTGCTTCTGTGCGTGTGTGAGGCGAGCGACCTTGGTGGC 120
DB 61 GAGGCTCAAAAGTTTGCTTCTGTGCGTGTGTGAGGCGAGCGACCTTGGTGGC 120
QY 121 CAGGAGTTGGGACGAAATTCAGGACCTCAAGAGACCTTCAAGGAGCTTCCCGGTGGA 180
DB 121 CAGGAGTTGGGACGAAATTCAGGACCTCAAGAGACCTTCAAGGAGCTTCCCGGTGGA 180
QY 181 GCGGGCTGTGCGGAGATCTGACCGGCTTCTCCAAAGCTTCTC----- 226
DB 181 GCGGGCTGTGCGGAGATCTGACCGGCTTCTCCAAAGCTTCTC----- 240
QY 227 -GATGACACACTGTTGGGACCGGTGGGCGGACGAGCCGGCTGGCGCGCTGCAGCT 285
DB 241 GGATGACACACTGTTGGGACCGGTGGGCGGACGAGCCGGCTGGCGCGCTGCAGCT 300
QY 286 GTTGGAAACCTATTCTCGGAGGCTCTGCGGAGCTGCGGAGCGGCTGGGAGCGCCGAC 345
DB 301 GTTGGAAACCTATTCTCGGAGGCTCTGCGGAGCTGCGGAGCGGCTGGGAGCGCCGAC 360
QY 346 GATCACTGGCTTCTTCGACACCGCAACCCCTGGACCTGGAGCCCGGCTGCCACCCGGCAG 405
DB 361 GATCACTGGCTTCTTCGACACCGCAACCCCTGGACCTGGAGCCCGGCTGCCACCCGGCAG 420
QY 406 CCGGTGATCTGCGCCACCCAGAGGAGAGAGCTTCTTCTGCGCTGCGGGCGGCTCTC 465
DB 421 CCGGTGATCTGCGCCACCCAGAGGAGAGAGCTTCTTCTGCGCTGCGGGCGGCTCTC 480
QY 466 CATCACAGCTGAGGGCTCAGAGCTCGGCTGCTGAGCGCTTCTGACCCAGGACAC 525
DB 481 CATCACAGCTTGGAGGCTCAGAGCTCGGCTGCTGAGCGCTTCTGACCCAGGACAC 540
QY 526 GCGGGATAGGCTTTTTCAGGGGCGAGCGCCAGGAGAGCTTGGAGCTGTGCTGCGGACCC 585

DB 541 GCGGATAGGCTTTTTCAGGCGCAGGCCAGGAGAGCCTGTGCTGTGCTGCGGACCC 600
QY 586 CTCAGGCTGGTGGTGGTGGAGAACGAGACCGGACAGACCGCTGGTTTCCAGCGCCCTA 645
DB 601 CTCAGGCTGGTGGTGGTGGAGAACGAGACCGGACAGACCGCTGGTTTCCAGCGCCCTA 660
QY 646 CCTGAGGAGGCGGCCCGCGGCCCAAGGCGGGAGGAGGCGGCTCCCTAGGAGGAGCGG 705
DB 661 CCTGAGGAGGCGGCCCGCGGCCCAAGGCGGGAGGAGGCGGCTCCCTAGGAGGAGCGG 720
QY 706 TCCCCAGTCTGTGCTTCCCGCGCTACGAGAGCAGCGCGCAGATGATGATGCTGCTGCC 765
DB 721 TCCCCAGTCTGTGCTTCCCGCGCTACGAGAGCAGCGCGCAGATGATGATGCTGCTGCC 780
QY 766 GCGGGGCGGCGCTGCGCGCTGTGGAACCTCAGACCGCGGCTGCTGCTATGCAGGTA 825
DB 781 GCGGGGCGGCGCTGCGCGCTGTGGAACCTCAGACCGCGGCTGCTGCTATGCAGGTA 840
QY 826 GCGGACCGGCGCGGCTTACTTCCCGCGGCTGCTGCGGCGGAGAGGCTGGGCGCTCT 885
DB 841 GCGGACCGGCGCGGCTTACTTCCCGCGGCTGCTGCGGCGGAGAGGCTGGGCGCTCT 900
QY 886 CCTCAGCGGAGCGGGTTCCTGTGAGGAGACGACCCCGCGGCTGAGGCCCGGGGCTTCCC 945
DB 901 CTTGAGCGGAGCGGGTTCCTGTGAGGAGACGACCCCGCGGCTGAGGCCCGGGGCTTCCC 960
QY 946 TGAACCTTCCAGGCGACCGGCTTCCCGCGGCTGCGGCGGAGAGCTTGGCGGGGCGC 1005
DB 961 TGAACCTTCCAGGCGACCGGCTTCCCGCGGCTGCGGCGGAGAGCTTGGCGGGGCGC 1020
QY 1006 CATCAGAGCGGCTGTGCGCTCAGCGAGGCGCTTGGAGGCGGCGCCACGCGCGCA 1065
DB 1021 CATCAGAGCGGCTGTGCGCTCAGCGAGGCGCTTGGAGGCGGCGCCACGCGCGCA 1080
QY 1066 GGGCGGCTCGAGGCTGCGTGGACTCTGTGCCGACCCACGAGCAGTGTGAGCGCG 1125
DB 1081 GGGCGGCTCGAGGCTGCGTGGACTCTGTGCCGACCCACGAGCAGTGTGAGCGCG 1140
QY 1126 AGGATCC 1132
DB 1141 AGGATCC 1147
RESULT 4
ADQ85428
ID ADQ85428 standard; cDNA; 1520 BP.
XX ADQ85428;
XX AC ADQ85428;
XX DT 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #2242.
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
OS WO2004060270-A2.
FN 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
PR (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
DR

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 2242; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
XX Sequence 1520 BP; 243 A; 512 C; 520 G; 245 T; 0 U; 0 Other;

Query Match 97.3%; Score 1101.8; DB 13; Length 1520;

Best Local Similarity 99.5%; Pred. No. 2.2e-207;

Matches 1127; Conservative 0; Mismatches 2; Indels 4; Gaps 2;

QY 1 AGCCATGGCAGGCCCCGATACCCAGTTCAGTGCAGGAGGCGCCCTGGTGAGATCAA 60
DB |||||||
QY 269 AGCCATGGCAGGCCCCCGATACCCAGTTCAGTGCAGGAGGCGCCCTGGTGAGATCAA 328
DB |||||||
QY 61 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCGTGGTGCAGAGCGGACACCTTCTGTGGC 120
DB |||||||
QY 329 GAGGCTCCAAACGTTTGGCTTCTCTGTGGCGTGGTGCAGAGCGGACACCTTCTGTGGC 388
DB |||||||
QY 121 CAGGAGTTGGAGCAATTCAGGAGCTCAAGAGACCTTCAAGAGACCTTCCCGGTGGA 180
DB |||||||
QY 389 CAGGAGTTGGAGCAATTCAGGAGCTCAAGAGACCTTCAAGAGACCTTCCCGGTGGA 445
DB |||||||
QY 181 GCGGCGCTGCTGCGAGATCTGACCGCTTCTCCCAAGAGCTTCTCGATGCACTGTGTT 240
DB |||||||
QY 446 GCGGCGCTGCTGCGAGATCTGACCGCTTCTCCCAAGAGCTTCTCGATGCACTGTGTT 505
DB |||||||
QY 241 GGGACCGCTGGGGCGCACAGCGCGCTGGCGCGCTTCTCGATGCACTGTGTTGAAACCTATT 300
DB |||||||
QY 506 GGGACCGCTGGGGCGCACAGCGCGCTGGCGCGCTTCTCGATGCACTGTGTTGAAACCTATT 565
DB |||||||
QY 301 TCGGAGGCTGCTGGCGACTGTCAGAGCGCGTGGCACCGAGCCCGACATCACTGGCTTCTTT 630
DB |||||||

DB 566 TCGGAGGCTGCTGGCGACTGACAGAGCGCGTGGCACGAGAGCCCGACGATCACTGGCTTCTT 625
QY |||||||
DB 361 CGCACCGCAACCCCTGAGACCTGAGAGCCCGCTGCCACCCGCGAGCCGCGGTGATCTCTGCC 420
DB |||||||
DB 626 CGCACCGCAACCCCTGAGACCTGAGAGCCCGCTGCCACCCGCGAGCGCGGTGATCTCTGCC 685
DB |||||||
QY 421 CACCCACAGAGGAGCAGCCTCTTTCTCGCGTGGCGGCCGCTCTTCCATCCACAGTCTGGA 480
DB |||||||
DB 686 CACCCACAGAGGAGCAGCCTCTTTCTCGCGTGGCGGCCGCTCTTCCATCCACAGTCTGGA 745
DB |||||||
QY 481 GGTCTAGAGCCTGCGCTGCTGAGCCCTCTGTACCCAGGACACCGGGATAGGCTTTT 540
DB |||||||
DB 746 GGTCTAGAGCCTGCGCTGCTGAGCCCTCTGTACCCAGGACACCGGGATAGGCTTTT 805
DB |||||||
QY 541 TCAGGCGCAGGCGCCAGAGAGCCTGAGACCTGCTGCTGCGCGCACCCCTCAGGCTTGGTGGCT 600
DB |||||||
DB 806 TCAGGCGCAGGCGCCAGAGAGCCTGAGACCTGCTGCTGCGCGCACCCCTCAGGCTTGGTGGCT 865
DB |||||||
QY 601 GGTGGAGAACGAGACCGGACAGCCCTGCTGTTTCCAGCGCCCTACTCTGAGAGAGCGGC 660
DB |||||||
DB 866 GGTGGAGAACGAGACCGGACAGCCCTGCTGTTTCCAGCGCCCTACTCTGAGAGAGCGGC 925
DB |||||||
QY 661 CCGGCGCCAAAGCGCGGAGGAGCCGCTCCCTAGGAGCAGCGGTCCCACTCTCTGTGC 720
DB |||||||
DB 926 CCGGCGCCAAAGCGCGGAGGAGCCGCTCCCTAGGAGCAGCGGTCCCACTCTCTGTGC 985
DB |||||||
QY 721 TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCGTCGCCGCGGCGGCGCGCT 780
DB |||||||
DB 986 TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCGTCGCCGCGGCGGCGCGCT 1045
DB |||||||
QY 781 GCGCGTGTGGAAACGTCAGACCGCGGCTGGTGGCTATCAGGTACG-CGCAACCGGCGCG 839
DB |||||||
DB 1046 GCGCGTGTGGAAACGTCAGACCGCGGCTGGTGGCTATCAGGTACGCGGAGCGGCGCG 1105
DB |||||||
QY 840 GCTACTTCCCGCGGTGCTGCTGCGGCCGAGAGGCTGCGCGCTCTCTCTGAGGAGGACGG 899
DB |||||||
DB 1106 GCTACTTCCCGCGGTGCTGCTGCGGCCGAGAGGCTGCGCGCTCTCTCTGAGGAGGACGG 1165
DB |||||||
QY 900 GGTTCGCTGGAGAGACGACCCCGCGGCTGAGCGCCGCGGCTTCCCTGAACCTCTCCAGG 959
DB |||||||
DB 1166 GGTTCGCTGGAGAGAGACCCCGCGGCTGAGCGCGGCTTCCCTGAACCTCTCCAGG 1225
DB |||||||
QY 960 CCACGCGCCTTCCCGCCACCGTGCACCCGACCTTTCGCGGCGGCGCATCCAGAGCGCT 1019
DB |||||||
DB 1226 CCACGCGCCTTCCCGCCACCGTGCACCCGACCTTTCGCGGCGGCGCATCCAGAGCGCT 1285
DB |||||||
QY 1020 GCTGACCGCTCACAGCAGGCGCCTGGAGCGGCGCCACGCGGCCAGGCGCGCTCGAG 1079
DB |||||||
DB 1286 GCTGACCGCTCACAGCAGGCGCCTGGAGCGGCGCCACGCGGCCAGGCGCGCTCGAG 1345
DB |||||||
QY 1080 GGTGCGTGACTCTGTGCGCGCACCCCGACGAGCAGTGAGCGCGAGGATCC 1132
DB |||||||
DB 1346 GGTGCGTGACTCTGTGCGCGCACCCCGACGAGCAGTGAGCGCGAGGATCC 1398
DB |||||||

RESULT 5

ADJ57074

ID ADJ57074 standard; cDNA; 1144 BP.

XX ADJ57074;

AC ADJ57074;

XX 06-MAY-2004 (first entry)

DT 06-MAY-2004 (first entry)

XX Coding sequence for p41Nox, a regulatory protein for Nox enzyme.

DE Nox regulatory protein; p41Nox; NADH oxidase; cytosolic; chromosome 16;

XX human; gene; ss.

KW Homo sapiens.

XX Key

Location/Qualifiers

5..1132

FT CDS

/*tag= a

FT

```

/ product= "Human p41Nox"
FO 2004007689-A2.
PN 22-JAN-2004.
XX 16-JUL-2003; 2003WO-US022246.
XX 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX (UYEM-) UNIV EMORY.
PA Lambeth JD, Cheng G;
XX WPI; 2004-122927/12.
PI P-PSDB; ADJ57075.
XX
XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
PT phosphate oxidase, useful for preparing a medicament for affecting
PT superoxide formation or regulating cellular proliferation in an animal or
PT a human.
XX
XX Claim 8; SEQ ID NO 7; 70pp; English.
XX
XX The present sequence is that of cDNA encoding a regulatory protein,
CC denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein sequence of
CC p41Nox was used as query in database screenings, and a genomic sequence
CC was identified from human chromosome 16. PCR primers based on this
CC sequence were used in RT-PCR experiments to look for expression in a
CC series of normal and cancer cell lines. A 350 bp PCR product was detected
CC in RNA from human foetal kidney, liver, lung and T84 colon cancer cells.
CC RNA from T84 cells and human foetal liver was then used to amplify the
CC full-length p41Nox cDNA by PCR. 13 Unique clones were obtained.
CC Sequencing revealed 4 p41Nox splice variants, of which nucleotide and
CC protein sequences are provided ADJ57068-ADJ57075. p41Nox proteins
CC function in the regulation of cell growth and are therefore implicated in
CC diseases involving abnormal cell growth, such as cancer. They may also
CC function in innate immune mechanisms of epithelial tissue or other
CC barrier cells, and hence may be involved in diseases of diminished
CC ability to fight infections or inflammatory conditions. The p41Nox
CC regulatory proteins and nucleic acids are useful for preparing a
CC medicament for affecting superoxide formation or regulating cellular
CC proliferation in an animal or a human. The regulatory proteins are also
CC useful in drug development, e.g. screening for drugs that regulate the
CC biological activity of the proteins, or in assays that relate to
CC assessment of abnormal growth or cellular proliferation including cancer.
XX
XX Sequence 1144 BP; 173 A; 404 C; 396 G; 171 T; 0 U; 0 Other;
SQ
Query Match 96.4%; Score 1091; DB 12; Length 1144;
Best Local Similarity 98.4%; Pred. No. 2.9e-205;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;
QY 1 AGCCATGCGAGCGCCCGATACCCAGTTTCAGTGCAGAGGGGAGCCCTTGGTGACAGTCAA 60
DB 1 AGCCATGCGAGCGCCCGATACCCAGTTTCAGTGCAGAGGGGAGCCCTTGGTGACAGTCAA 60
QY 61 GAGGCTCCAAACGTTTGCTTCTCTGTGCGTGTGTCAGAGCGGAGCGACACTTTCGTGCG 120
DB 61 GAGGCTCCAAACGTTTGCTTCTCTGTGCGTGTGTCAGAGCGGAGCGACACTTTCGTGCG 120
QY 121 CAGGAGTTGGACGAGTTCAGGACAGCTCAAGAACCCCTCAAGAGACCTTCCGGTGGGA 180
DB 121 CAGGAGTTGGACGAGTTCAGGACAGCTCAAGAACCCCTCAAGAGACCTTCCGGTGGGA 177
QY 181 GCGCGGCTGCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTC----- 226
DB 178 GCGCGGCTGCTGCGGAGATCTGACCGGTTCTCCAAAGCTTCTCGGTACGGCAGCGCT 237
QY 227 -GATGCAACCACTGTTGGAGCCGCTGGGGCGACACGAGCGCGCGCTGGCGCGCTGCACT 285
DB 238 GATGCAACCACTGTTGGAGCCGCTGGGGCGACACGAGCGCGCGCTGGCGCGCTGCACT 297
```

```

QY 286 GTTGGAAACCTATTCTCGAGGCTGCTGGCGACTGACAGCGCGTGGCAACGAGCGCCGAC 345
DB 298 GTTGGAAACCTATTCTCGAGGCTGCTGGCGACTGACAGCGCGTGGCAACGAGCGCCGAC 357
QY 346 GATCACTGGCTTCTTTCGCAACCCCTCGAACCTGAGAGCCCGCTGCCACACCGGCGAG 405
DB 358 GATCACTGGCTTCTTTCGCAACCCCTCGAACCTGAGAGCCCGCTGCCACACCGGCGAG 417
QY 406 CCGGCTGATCTGCTCCACCCCGAGAGGAGACGCTTCTTCTCGCGCTGGGGCGGCTCTC 465
DB 418 CCGGCTGATCTGCTCCACCCCGAGAGGAGACGCTTCTTCTCGCGCTGGGGCGGCTCTC 477
QY 466 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGAGCCCTTCTGTACCCAGGACAC 525
DB 478 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGAGCCCTTCTGTATCCAGGACAC 537
QY 526 GCGGGATAGGCTTTTTCAGGCGCAGGCGCCAGAGAGCTGAGACCTGCTGCTGCGGACACC 585
DB 538 GCGGGATAGGCTTTTTCAGGCGCAGGCGCCAGAGAGCTGAGACCTGCTGCTGCGGACACC 597
QY 586 CTCAGGCTGCTGCTGCTGAGAGAACGAAAGACCCGAGACCGCTGCTGCTTCCAGGCGCTTA 645
DB 598 CTCAGGCTGCTGCTGCTGAGAGAACGAAAGACCCGAGACCGCTGCTGCTTCCAGGCGCTTA 657
QY 646 CCTGGAGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCGCTCCCTAGGAGGACGCGG 705
DB 658 CCTGGAGAGCGCGCCCGCGGCCAAGCGCGGAGGAGGCGCTCCCTAGGAGGAGCGG 717
QY 706 TCCCGAGTCTGCTTCTCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTGCTGCTGCC 765
DB 718 TCCCGAGTCTGCTTCTCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTGCTGCC 777
QY 766 GCGGGGCGCGCTGCTGCTGCTGTTGAAACGTCAGACCGCGCTGCTGCTGCTATGAGGTA 825
DB 778 GCGGGGCGCGCTGCTGCTGCTGTTGAAACGTCAGACCGCGCTGCTGCTATGAGGTA 837
QY 826 GCGGAGCGCGCGCGCTTCTCCCGCGCTGCTGCTGCGCGCGGAGAGGCTGGCGCTCT 885
DB 838 GCGGAGCGCGCGCGCTTCTCCCGCGCTGCTGCTGCGCGCGGAGAGGCTGGCGCTCT 897
QY 886 CCTGAGCGGAGCGGGTTCGCTGAGAGAGACGACCCCGCGGCTGAGGCGCGGGCTTCCC 945
DB 898 CCTGAGCGGAGCGGGTTCGCTGAGAGAGACGACCCCGGGTTCAGGCGCGGGCTTCCC 957
QY 946 TGAACCTTCCAGGCGACCGCGCTTCCCGCGCTGCTGCTGCGCGCGGAGGCTTCCGCGGCG 1005
DB 958 TGAACCTTCCAGGCGACCGCGCTTCCCGCGCTGCTGCTGCGCGCGGAGGCTTCCGCGGCG 1017
QY 1006 CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
DB 1018 CATCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 1066 GGGCGCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
DB 1078 GGGCGCGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1137
QY 1126 AGGATCC 1132
DB 1138 AGGATCC 1144
```

```

RESULT 6
AAS89078
ID AAS89078 standard; cDNA; 1911 BP.
XX
AC AAS89078;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #24882.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
```

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US0008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX PA Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR P-PSDB; ABG24891.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 24882; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1911 BP; 397 A; 545 C; 587 G; 382 T; 0 U; 0 Other;
Query Match 70.5%; Score 798; DB 5; Length 1911;
Best Local Similarity 98.2%; Pred. No. 1.2e-147;
Matches 823; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
QY 5 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGAGCCCTGTGCAGATCAAGAGG 64
DB 1 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGAGCCCTGTGCAGATCAAGAGG 60
QY 65 CTCCTCAAAAGTTTGCCTTCTGTGTGCGCTGTGCAGAGCGGAGGACACCTTGTGCGCAGG 124
DB 61 CTCCTCAAAAGTTTGCCTTCTGTGTGCGCTGTGCAGAGCGGAGGACACCTTGTGCGCAGG 120
QY 125 AGTTGGGAGGATTCAGGAGGCTCAAGAGAGCCCTCAAGGAGACCTTCCCGGTGGAGGCG 184
DB 121 AGTTGGGAGGATTCAGGAGGCTCAAGAGAGCCCTCAAGGAGACCTTCCCGGTGGAGGCG 180
QY 185 GGCCTGCTCGGAGATCTGACCGCGTTCCTCCCAAGCTTCTC-----GAT 229
DB 181 GGCCTGCTCGGAGATCTGACCGGTTCCTCCCAAGCTTCTCGGTCAAGCCAGCCTGGAT 240
QY 230 GCACCACTTTGGAGCGCTGGGGCGCAGAGCCGCGCTGCGCGCTGAGCTGTG 289

DB 241 GCACCACTTTGGAGCGCTGGGGCGCAGAGCCGCGCTGCGCGCTGCGAGCTGTG 300
QY 290 GAAACCTATTTCGGAGGCTGTGCGGACTGACAGAGCGGTGCGACGAGCCGAGATC 349
DB 301 GAAACCTATTTCGGAGGCTGTGCGGACTGACAGAGCGGTGCGACGAGCCGAGATC 360
QY 350 ACTGGCTTTCTGCGACCCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCCGCAGCCGG 409
DB 361 ACTGGCTTTCTGCGACCCGCAACCCCTGGACCTGGAGCCCGCGCTGCCACCCCGCAGCCGG 420
QY 410 GTGATCTGCGCCACCCCGCAGAGGAGCGCTCTTTCTGCGCGCTGCGGCGCGCTTCTCATC 469
DB 421 GTGATCTGCGCCACCCCGCAGAGGAGCGCTCTTTCTGCGCGCTGCGGCGCGCTTCTCATC 480
QY 470 CACAGTCTGAGGCTCAGAGCGCTGCGCTGCGAGCCCTTCTGTATCCAGGACACGCGG 529
DB 481 CACAGTCTGAGGCTCAGAGCGCTGCGCTGCGAGCCCTTCTGTATCCAGGACACGCGG 540
QY 530 GATAGCGCTTTTCAGGCGCAGGCGCCAGGAGAGCGCTGGACGTGCTGCTGCGGACCCCTCA 589
DB 541 GATAGCGCTTTTCAGGCGCAGGCGCCAGGAGAGCGCTGGAGCGTGTGCTGCGGACCCCTCA 600
QY 590 GCGTGTGCTGCTGGAGAACGAGACCGGAGAGCCGCGCTGCTTTCAGGCGCGCTTACCTG 649
DB 601 GCGTGTGCTGCTGGAGAACGAGACCGGAGAGCCGCGCTGCTTTCAGGCGCGCTTACCTG 660
QY 650 GAGGAGCGCGCCCGGCGCAAGCGCGGAGGAGCGCCCTCCTAGGAGAGCGGTGCC 709
DB 661 GAGGAGCGCGCCCGGCGCAAGCGCGGAGGAGCGCCCTCCTAGGAGAGCGGTGCC 720
QY 710 CAGTTCTGTGCTTCCCGCGCTTACGAGACGACCGCGCAGATGAGCTGTGCTGCGCGCG 769
DB 721 CAGTTCTGTGCTTCCCGCGCTTACGAGAGCAGCGCGCAGATGAGCTGTGCTGCGCGCG 780
QY 770 GGGGCGCGCGCGCGGTGTGGAACCTCAGACCGCGCGCTGCTGCTATGCAAGTACG 827
DB 781 GGGGCGCGCGCGGTGTGGAACCTCAGACCGCGCGCTGCTGCTATGCAAGTACG 838
RESULT 7
AAS89434
ID AAS89434 standard; cDNA; 2881 BP.
XX AAS89434;
AC AAS89434;
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #25238.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US0008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX PA Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR P-PSDB; ABG25247.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
 PS Claim 1; SEQ ID NO 25238; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2881 BP; 643 A; 803 C; 810 G; 625 T; 0 U; 0 Other;
 Query Match 70.5%; Score 798; DB 5; Length 2881;
 Best Local Similarity 98.2%; Pred. No. 1.2e-147;
 Matches 823; Conservative 0; Mismatches 0; Indels 15; Gaps 1;
 5 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGAGCCCTGTCAGATCAGAGG 64
 Db 1 ATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGAGCCCTGTCAGATCAGAGG 60
 65 CTCGAAACGTTTGCCCTTCTGTGCGCTGTTGAGACGCGACACCTTGTGCGCAGG 124
 Db 61 CTCGAAACGTTTGCCCTTCTGTGCGCTGTTGAGACGCGACACCTTGTGCGCAGG 120
 125 AGTTGGGACGAATTCAGGCAGCTCAAGAGAGCCCTCAAGGAGACCTTCCCGTGGAGCG 184
 Db 121 AGTTGGGACGAATTCAGGCAGCTCAAGAGAGCCCTCAAGGAGACCTTCCCGTGGAGCG 180
 185 GGCCTGCTCGGAGATCTGACCGGTTTCTCCAAAGCTTCTC-----GAT 229
 Db 181 GGCCTGCTCGGAGATCTGACCGGTTTCTCCAAAGCTTCTCGGTGAGGCCAGCTGGAT 240
 230 GCACCACTGTTGGGACGCTGGGGCGCACGAGCCGCGCTGGCGCGCTGCGAGCTGTG 289
 Db 241 GCACCACTGTTGGGACGCTGGGGCGCACGAGCCGCGCTGGCGCGCTGCGAGCTGTG 300
 290 GAAACCTATTCTCGGAGGCTGTGCGGACTGCGAGAGCGCTGGCAGGAGCCGACGATC 349
 Db 301 GAAACCTATTCTCGGAGGCTGTGCGGACTGCGAGAGCGCTGGCAGGAGCCGACGATC 360
 350 ACTGCTTCTTCGACCCGACCCCTGGAACCTGAGGCCGCGCTGCGACCCGCGAGCCGG 409
 Db 361 ACTGCTTCTTCGACCCGACCCCTGGAACCTGAGGCCGCGCTGCGACCCGCGAGCCGG 420
 410 GTGATCTCGCCACCCAGAGAGAGCAGCTCTTTCTCGCGCTGGGGCGCGCTCTCCATC 469
 Db 421 GTGATCTCGCCACCCAGAGAGAGCAGCTCTTTCTCGCGCTGGGGCGCGCTCTCCATC 480
 470 CACAGTCTGGAGGCTCAGAGCTGCGCTGCGAGCCCTTCTGTACCCAGAGACGCGG 529
 Db 481 CACAGTCTGGAGGCTCAGAGCTGCGCTGCGAGCCCTTCTGTACCCAGAGACGCGG 540
 530 GATAGGCCCTTTTCAGGCGCAGCCAGAGAGCCCTGAGAGCTGCTGCGGACACCCCTCA 589
 Db 541 GATAGGCCCTTTTCAGGCGCAGCCAGAGAGCCCTGAGAGCTGCTGCGGACACCCCTCA 600
 590 GGCTGGTGGCTGGAGAAAGAGAGCCGAGAGCCGCGCTGGTTTCCAGCGCCCTACCTG 649

Db 601 GGCTGGTGGCTGGAGAAAGAGAGCCGCGAGAGCCCTGGTTTCCAGCGCCCTACCTG 660
 Qy 650 GAGAGAGCGCGCCCGCGGCCAAGCGCGGAGAGAGCCCGTCCCTAGGAGAGCGGTCCC 709
 Db 661 GAGAGAGCGCGCCCGCGGCCAAGCGCGGAGAGAGCCCGTCCCTAGGAGAGCGGTCCC 720
 Qy 710 CAGTTCTGTGCTTCCCGCGCTACGAGAGAGCGCGCGAGATGAGCTGCTCCGTGCCCGG 769
 Db 721 CAGTTCTGTGCTTCCCGCGCTACGAGAGAGCGCGCGAGATGAGCTGCTCCGTGCCCGG 780
 Qy 770 GGGCGCGCGTGGCGCGCTTTCGAAACGTCAGACCGCGCTGCTGGCTATGCAAGTACG 827
 Db 781 GGGCGCGCGTGGCGCGCTTTCGAAACGTCAGACCGCGCTGCTGGCTATGCAAGTACG 838
 RESULT 8
 ADM02780/c
 ID ADM02780 standard; cDNA; 2922 BP.
 XX
 AC ADM02780;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human cDNA of the invention SEQ ID NO:1465.
 XX
 KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
 XX
 OS Homo sapiens.
 XX
 PN EPI347046-A1.
 XX
 PD 24-SEP-2003.
 XX
 PF 12-APR-2002; 2002EP-00008400.
 XX
 PR 22-MAR-2002; 2002JP-00137785.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-723558/69.
 DR P-PSDB; ADM05223.
 XX
 PT New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX
 PS Claim 1; SEQ ID NO 1465; 305pp; English.
 XX
 The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.
 XX
 SQ Sequence 2922 BP; 492 A; 954 C; 980 G; 496 T; 0 U; 0 Other;
 Query Match 30.2%; Score 342.4; DB 11; Length 2922;
 Best Local Similarity 84.9%; Pred. No. 4.8e-58;
 Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;
 704 GGTCCCAGTTCTGTGTTTCCCGCGCTACGAGAGAGCCGCGAGATGAGCTGTCGCTG 763
 Db 679 GGTCCCAGTTCTGTGTTTCCCGCGCTACGAGAGAGCCGCGAGATGAGCTGTCGCTG 620

Qy	764	CCCGCGGGGCGCGGTGCGCGTCTGGAAACGTACAGCCGCGGCTGGTGGCTAT-----	818
Db	619		818
Qy	819	-----	818
Db	559	TACGCGGAGCGGCGTGGGACGGGCAGGCTAGCCGAGCGCGGCGGGCCCTAAACACAC	500
Qy	819	-----GACGGTACGGCGACCGGGCGGGCTACTCCCGCGGTGCTGCTCGGGCCG	868
Db	499	CCCGCCCTCGCAGGTACGGCGACCGGGCGGCTACTCCCGCGGTGCTGCTGGGCGG	440
Qy	869	GAAGGGTGGGCGTCTCTGAGCGGGACGGGGTTCGTGAGGAGACGACCCCGCGGGT	928
Db	439	GAAGGGTGGGCGTCTCTGAGCGGACGGGGTTCGTGAGGAGACGACCCCGCGGGT	380
Qy	929	GAGGCCCGGGCTTCCCTGAACTCCAGGCCACCGCCCTCCCGCCAGCGTCCCGACCC	988
Db	379	GAGGCCCGGGCTTCCCTGAACTCCAGGCCACCGCCCTCCCGCCAGCGTCCCGACCC	320
Qy	989	CGACTTTCGCGCGGCGCATCCAGAGCGGTGCTGCACCGTCAACGAGGGCCCTGGAG	1048
Db	319	CGACTTTCGCGCGGCGCATCCAGAGCGGTGCTGCACCGTCAACGAGGGCCCTGGAG	260
Qy	1049	CGGCGCCACGCGGCCAGGGCCGCTTCGAGGGTCGTGGACTCTGTGCGCACCCCGACG	1108
Db	259	CGGCGCCACGCGGCCAGGGCCGCTTCGAGGGTCGTGGAATCTGTGCGCACCCCGACG	200
Qy	1109	ACGAGCAGTGAGCGCAGGATCC	1132
Db	199	ACGAGCAGTGAGCGCAGGATCC	176

RESULT 9

AA53491
ID AAX53491 standard; DNA: 114955 BP.

AC AAX53491;

DT 05-JUL-1999 (first entry)

DE Human adenosine A1 receptor antisense oligonucleotide fragment.

Antisense oligonucleotide; multiple target; antisense treatment;
impaired respiration; inflammation; lung disease;
pulmonary vasoconstriction; inflammation; allergic rhinitis;
acute asthma; allergy; asthma; impaired respiration;
respiratory distress syndrome; pain; cystic fibrosis;
pulmonary hypertension; pulmonary vasoconstriction; emphysema;
chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
colon cancer; breast cancer; lung cancer; pancreatic cancer;
hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
prostate cancer; ss.

OS Synthetic.

PN WO9913886-A1.

25-MAR-1999.

XX 17-SEP-1998: 98WO-US019419.

XX
PR 17-SEP-1997: 97US-0059160P.

PR 09-JUN-1998; 9805-00093972.
XX

PA (UYEC-) UNIV EAST CAROLINA.
YY

PI Nyce JW;
vv

DR WPI; 1999-229400/19.

PT New antisense oligonucleotides used in treatment of, e.g. pulmonary

PT
XX vasoconstriction.
FS Disclosure; Page 37; 120pp; English.
XX

The specification describes antisense oligonucleotides (AAX52869-X55271) directed against at least 2 mRNAs selected from target genes, coding and non-coding regions of RNAs corresponding to target genes, gene initiation codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-end and the juxta-section between coding and non-coding regions and all segments of RNAs encoding proteins associated with one or more diseases, conditions or mixtures. The antisense oligonucleotides may be derived from sequences AAX5272-74. These multiple target oligonucleotides (specifically AAX53180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

CC Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;
XX SQ

Query Match	6.8†	Score	77.2;	DB 2;	Length	114955;
Best Local Similarity	33.4†	Pred. No.	7.9e-06;			
Matches	296;	Conservative	92;	Mismatches	484;	
				Indels	14	Gaps 4;

Qy	251	GGGCGCAGAGCGCGCGCTGGCGCGCTGCAGCTGTTGGAAA	CCTATCTCTCGAGGCTG	310
Db	104410	GGNNNDNCCGCBGCBGGGCGCGCGCGGCGGNNNDN	CCGCBGCGCBGGCGCG	104469
Qy	311	CTGCGCATCTGAGAGCGCTGGCAGGAGCCCGAC----	GATCACTGGCTTCTTCGCACC	366
Db	104470	CCGCGGCCNNNDNNDNCCGCBGGCCBGGCGCGCGCGGCS	NNNDNNDNCCGCBGCCBGG	104529
Qy	367	GCAACCCCTGGACCTGAGAGCCGCGCTGCCACCCGCGAG	CGGGTGATCTTCCCCACCCC	426
Db	104530	GCGCGCCGCGGNNNDNCCGCBGGCCBGGGCGCGCGGCS	NNNDNNDNCCGCBGGCCBGG	104589

QY	427	AGAGGAGCAGCTTTTCTTCGGCTGCGGC CGCCCTCTCCATCCACACTCTGAGGCTCA	486
DB	104590	GCGCGCGCGCCSNNNDNNC CGBGGC CBGGCGCGCGCSNNNDNNC GCBGGC CBGGC	104649

Qy	487	GAGCCTGGCTGCCTGAGCCCT-----TCGTACCCAGGACACGCGGGATAGGCCCTTT	540
DB	GGCGCGNNNNNNCCGCGGGCGGGCGCGCGNNNNNNCCGCGGGCGGGCGCGCGSNN	104709	104709

Qy	541	TACGGCGAGCCGAGAGCCTGGACGTCTGTCGGCACCCCCTACGGCTGGTGGCT	600
Dβ	104710	NDDNCCGCGGGCGCGGCGGNNNDNCCGCGGGCGCGGGCGCSNNNDNCCBGGGCC	104769

Qy 601 GGTGGAGAACAGACCGGCAGACCGCCTGTTCTCAGCGCCT--ACCTGGAGGAGCGC 658
Db 104770 BGGCGCGNNNDNNCCGCGGGCGGGCGNNNDNNCCGCGGGCGGGGNNNDNNCCGCGC 104829

QY 659 GCCCCGGSCCAAGGCCGSGAGGAGGCCCGTCCCTAGCGAGCAGCGGTCCGCCAGTTCCTCT 718

DB 104830 GCCBGGNNNNNNCCGCGGGCCBGGNNNDNNCCGCBGGCCBSNNNDNNCCGGGGCCSNN 104889

Qy 71.9 GCTTCCCGCGCTTACGAGAGCAGCCGCGCAGATGAGTGTCTCGTCCCGCGCGCGCGCGCGC 778
Db 104890 NDNNCGCGCCG 104949

Oy 779 GTGCGCGTGTGGAAACGTACAGCGGGCTGTGGCTATGCAAGTACGGCGACCGGG- 837
Db 104950 CCGGCCGGCCGNNNDNNCBGGCCBGGGCGCGCCGCGCGGCGNNNNNBGGCC 105009

Q7 838 - GGGGCTACTCCCGCGCGTGCTGCTGGGGCCGGAAGGCGTCTCTGTACGGGA 896
||||| | ||||| : | |||| | |||| | |||| | |||| :

D	b		133	GCCCACCGCCGCTGC	CCCCCAGCATGGAGCGGCCGCGGTGCGCCGGTCCCGCT	197		
Q	y		385	GCCGCGCTGCCACC	CGCGCAGCCGGTGATCTCTGCCACCCAGAGGACGACCTTTTC	444		
D	b		193	TCCGCTGCTGCTGC	CGCGCGCTTTCGCTGCGCGCGGAGTGGACGCGGATGCTCT	252		
Q	y		445	TGCGCTGCGGCGCG	CCTCTTCATCAAGTCTGGAGGCTTAGAGCTTGGCTGCTGCTGCA	504		
D	b		253	CCTGGAGGCTGCTGT	GCGGACCGGATGCGCTCATCAAGAAGACTGCTCGCT	312		
Q	y		505	GCCCTTCTGTACCC	AGCACCGCGGATAGGCTTTTCAGGCGCAGGCCAGGAGGCT	564		
D	b		313	GCCATATGTCAGGA	TCCAAGATGCAAGGATGTCAGGAGCAGTCTGCCACAGCA	372		
Q	y		565	GGACGTGCTGCTGC	CGCGCACCCCTCAGGCTGTGGCTGGAGAACGAAAGCCG	618		
D	b		373	GCTGGAGGAGCTGC	ACTGTGCCACGGCATCAGCCTTGCCACGAGCAGGACCG	426		
RESULT 12								
A	B	V	23338	standard; cDNA; 2781 BP.				
A	C	A	ABV23338;					
X	X		16-SEP-2002	(first entry)				
D	E		Human prostate expression marker cDNA 23329.					
X	X		Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;					
K	M		pharmacogenomic marker; gene; ss.					
X	X		Homo sapiens.					
O	S		WO200160860-A2.					
P	N		23-AUG-2001.					
P	F		20-FEB-2001; 2001WO-US005171.					
X	X		17-FEB-2000; 2000US-0183319P.					
P	R		16-MAR-2000; 2000US-0189862P.					
P	R		25-MAY-2000; 2000US-0207454P.					
P	R		09-JUN-2000; 2000US-0211314P.					
P	R		18-JUL-2000; 2000US-0219007P.					
P	R		13-DEC-2000; 2000US-0255281P.					
X	X		(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.					
P	A		Schlegel R, Endege WO, Monahan JE;					
X	X		WPI; 2001-662795/76.					
Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.								
P	S		Claim 1; Page 6215; 11750pp; English.					
C	C		The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:					
C	C		(a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker					
X	X		Sequence 2781 BP; 569 A; 803 C; 825 G; 577 T; 0 U; 7 Other;					
Q	U		Query Match 5.0%; Score 56.8; DB 5; Length 2781;					
B	e		Best Local Similarity 47.8%; Pred. No. 0.069;					
M	a		Matches 198; Conservative 0; Mismatches 212; Indels 4; Gaps 1;					
Q	y		205	CGCGCTTCTCCCAA	AGCTTCGATGCCACCTATTCTCGAGGCTGTGGCGACGCGG	264		
D	b		17	CGGGTTCCTCCCG	GGGCTACGCGCGGCTTGGCCAGCGTGTGCTGCCGAGCTGCGCG	76		
Q	y		265	CGCGCTTGGCGCG	CCTCGACTGTGGAACTATTCTCGAGGCTGTGGCGACTGCGAGA	324		
D	b		77	GAGCTGTGGAGCC	CGCGCGCTGCCCCAGAC-----CGCGCGCGCGCTTTGTCGCGCGC	132		
Q	y		325	GGCGGTGGCAGG	AGCCCGACGATCACTGGCTTTCTTCGACGCGCAACCCCTGGACCTGGA	384		

Query Match	5.0%;	Score 56.8;	DB 5;	Length 2781;
Best Local Similarity	47.8%;	Pred. No. 0.069;		
Matches 198;	Conservative	0;	Mismatches 212;	Indels 4;
Gaps 1				
Qy	205	CCGCGTTCTCCCAAAGCTTCTCGATGCACCACTGTTGGAGCGCGTGGGGCGGCACGAGCCG	264	
Db	17	CGGTTCTCCCGGGCTACGCGCGGCCCTGCCACAGTTGGCTGCCGAGGCTCGGCCG	76	
Qy	265	CGGCTGGCGGCGCTGCGAGCTGTTGGAAACCTATTCTCGGAGGCTGTGGGACATGCAGA	324	
Db	77	GAGCGTGGAGCCCGCGCGCTGCCCCAGGAC-----CGCGCCGCGGCCCTTTGTCGCCGCC	132	
Qy	325	GCGCGTGGCAGGAGCCCGACGATCACTGGCTTCTTCGCACCGCAACCCCTCGAGACTCGGA	384	
Db	133	GCCACCGCGTGCCTCGCCGCGCCCATGAGCGCGCGCGCTGCGCGCGGGTCCCGCT	192	
Qy	385	GCCCGCGCTGCCACCGGCAGCCGGGTGATCTGCGCCACCCGAGGAGAGCCCTCTTTC	444	
Db	193	TCCGCTGCTGTGCTCGGCGCCCTTGCCTGCTGCGCGCGGAGTGGACGGGATGTCTCT	252	
Qy	445	TCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTTCGCTGCTCGCA	504	
Db	253	CCTGGAGGCTGCTGTCGGACGGACACCGGATGGCCACTCATCAGAAGGACTGCTCGCT	312	
Qy	505	GCCCTTCTGTACCCAGGACACGCGGGATAGCCCTTTTCAGGCGGAGGCCACGAGGAGCCT	564	
Db	313	GCCATATGCTACGGAATCCAAAGAAATGCAGGATGGTGCAGGAGCATGCTGCCACAGCCA	372	
Qy	565	GGACGTGCTGTCGGCACCCCTCAGGCTGGTGGCTGGTGGAGAACGAGACCG	618	
Db	373	GCTGGAGGAGCTGCACTGTGTCACGGGCATCAGCTGCGCCCAACGACGAGGACCG	426	

RESULT 13	
ABD33012	
ID	ABD33012 standard; cDNA; 1850 BP.
XX	
AC	ABD33012;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Human cancer-associated cDNA HR21-032.3.
XX	
KW	Human; ss; cancer-associated protein; gene; cytostatic; cancer;
KW	leukaemia; lymphoma; CAP.
XX	
OS	Homo sapiens.
XX	
PN	WO2004074320-A2.
XX	
PD	02-SEP-2004.
XX	
PF	17-FEB-2004; 2004WO-US004730.
XX	
PR	14-FEB-2003; 2003US-00367094.
PR	14-MAR-2003; 2003US-00388338.
PR	13-APR-2003; 2003US-00417375.
PR	13-JUN-2003; 2003US-00461862.
PR	15-SEP-2003; 2003US-00663431.
PR	15-DEC-2003; 2003US-00737318.

claim 1; seqid 787; 310pp; English.

The invention relates to an isolated nucleic acid comprising at least 10 contiguous nucleotides of any of the 233 polynucleotide sequences given in the specification, or its complement. The nucleic acids encode cancer-associated proteins. Also included are an expression vector comprising the isolated nucleic acid cited above, a host cell comprising the above recombinant nucleic acid or expression vector, a microarray for detecting a cancer-associated (CA) nucleic acid comprising at least one probe comprising at least 10 contiguous nucleotides of any of the above-mentioned nucleotide sequences, an isolated polypeptide (encoded within an open reading frame of a CA sequence selected from any of the 95 polynucleotide sequences as mentioned in the specification, or its complement), an isolated antibody, (for its antigen binding fragment) that binds to the above polypeptide, a hybridoma that produces the above monoclonal antibody, a pharmaceutical composition comprising the above antibody and a pharmaceutical excipient, a kit for detecting cancer cells (comprising the antibody cited above, methods for diagnosing cancer or for detecting the presence or absence of cancer cells in an individual, a method for inhibiting growth of cancer cells in an individual, a method for delivering a therapeutic agent to cancer cells in an individual, an electronic library comprising the above polynucleotide or polypeptide (or their fragments), methods of screening for anticancer activity or for a bioactive agent capable of modulating the activity of a CA protein (CAP), methods for detecting cancer associated with expression of a polypeptide in a test cell sample, a method for treating cancers and a method for inhibiting the expression of a CA gene in a cell. The composition and methods are useful for detecting, diagnosing, preventing and treating cancers, especially lymphoma and leukaemia. These may also be used in screening for agents that modulate cancer. The present sequence is a human CAP cDNA sequence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct sequences

Seq	Sequence	1850 BP	358 A	559 C	575 G	358 T	0 U	0 Other	
	Query Match	4.9%	Score	55.6	DB	13	Length	1850	
	Best Local Similarity	49.3%	Pred. No.	0.12					
	Matches	145	Conservative	0	Mismatches	149	Indels	0	
	Gaps	0							
Qy	325	GC	CGTGGC	ACGGAG	CCGACG	ATCATCTT	TCGCA	CGCAACCCCTGGACCTGGA	384
Db	83	GCC	ACCGCC	GTCCG	CGCCAT	GTGAG	CGCG	CGCGCTCCGCT	142
Qy	385	GCC	CGCGCT	GCCAC	CCGGCAGC	CGGGTGT	ATCCT	GCCACCCGAGAGGAGCAGCCTCTTTC	444
Db	143	TCC	GCTGTGTG	TCGCGG	CCCTT	GC	GCTG	CGCGCGAGTGGAGCGCGATGTCT	202
Qy	445	TCG	CGCTGGG	CGCCCT	CTCTCA	TCCA	CAGTCTG	AGGCTCAGAGCCTGCGCTGCTGCA	504
Db	203	CCT	GGAGCCTG	CTGTG	CGGAC	CGGAC	ACCGATGG	CCACTCATCAAGAGACTGCTCGCT	262
Qy	505	GCC	CTTCTGT	TACCCAGG	ACACG	CGGATAG	GCCCTTTT	CAGCGCGCAGGCGCCAGGAGAGCCT	564
Db	263	GCA	TATGTCTAC	GGAATCC	AAAGAT	GCAGAG	TGGTGC	AGGACAGTGTGCCACAGCCA	322
Qy	565	GG	ACGTGCTGT	CGGCA	CC	CCCTC	AGGCTGTG	TGCTGGTGAGAACGAAAGACCG	618
Db	323	GCT	GGAGAGCT	GACTGTG	TGCACGGG	CATC	CAGCTGT	GCCCAACGAGCAGGACCG	376

RESULT 14
ACN42908
ID ACN42908 standard; cDNA; 2196 BP.
XX
XX
AC ACN42908;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:1783.
DE
XX
XX ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW

XX dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX 12-SEP-2003; 2003WO-US028227.
XX PF 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCYTE) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F, Hartshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV, Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP, Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH, Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL, Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES, Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D, Patury S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
XX P-PSDB; ABM94256.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.
XX Claim 1; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorders, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp polynucleotide of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 2196 BP; 415 A; 708 C; 672 G; 401 T; 0 U; 0 Other;
Query Match 4.9%; Score 55.6; DB 13; Length 2196;
Best Local Similarity 49.3%; Pred. No. 0.12;
Matches 145; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 325 GCGCGTGGCAGCGAGCCGCGGATGATCTTTCGACCCGACACCCCTGACCTGGA 384
DB 308 GCCACCGCCGCTGCGCCGCGCCATGAGCGCGCGCGCGGATGAGCGGATGCGCT 367
QY 385 GCGCGCTGCCACCGCGAGCGCGGATGATCTTTCGACCCGACAGAGGAGCGCTCTTTC 444
DB 368 TCGCTGCTGCTGCTGCGCGCTTTCGCTGCTGCGCGCGGATGAGCGGATGCTCT 427
QY 445 TCGCGCTGCGCGCGCTTTCATCATCAGTCTGAGGCTGAGCGCTGCGCTGCTGCA 504
DB 428 CTGAGGCGCTGCTGTCGCGACGACCGGATGCGGCTCATCATGAGGAGCTGCTGCT 487
QY 505 GCGCTTCTACCGAGCAGCGGATAGGCTTTTCAGGCGCAGGCGCAGGAGCT 564
DB 488 GCCATATGCTACGAATCCAAAGATGACGAGTGGTGACGAGGAGCTGTCGCACAGCCA 547

QY 565 GGACGTGCTGCTGCGGACCCCTCAGGCTGGTGGTGGAGAACGAGACCG 618
DB 548 GCTGGAGGAGCTGCACTGTGCGCACGCGCATCAGCTGGCCACGAGCAGACCG 601
RESULT 15
ABD33013
ID ABD33013 standard; cDNA; 2215 BP.
XX ABD33013;
AC ABD33013;
XX 18-NOV-2004 (first entry)
XX Human cancer-associated cDNA HR21-032.4.
XX Human; ss; cancer-associated protein; gene; cytostatic; cancer;
XX leukemia; lymphoma; CAP.
XX Homo sapiens.
XX WO2004074320-A2.
XX 02-SEP-2004.
XX 17-FEB-2004; 2004WO-US004730.
XX 14-FEB-2003; 2003US-00367094.
XX 14-MAR-2003; 2003US-00388838.
XX 15-APR-2003; 2003US-00417375.
XX 13-JUN-2003; 2003US-00461862.
XX 15-SEP-2003; 2003US-00663431.
XX 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX P-PSDB; ABO84711.
XX New isolated cancer-associated polynucleotides and polypeptides useful for diagnosing, preventing or treating cancers, especially lymphoma and leukemia, or in screening for agents that modulate cancer.
XX claim 1; seqid 789; 310pp; English.
XX The invention relates to an isolated nucleic acid comprising at least 10 contiguous nucleotides of any of the 233 polynucleotide sequences given in the specification, or its complement. The nucleic acids encode cancer-associated proteins. Also included are an expression vector comprising the isolated nucleic acid cited above, a host cell comprising the above recombinant nucleic acid or expression vector, a microarray for detecting a cancer-associated (CA) nucleic acid comprising at least one probe comprising at least 10 contiguous nucleotides of any of the above-mentioned nucleotide sequences, an isolated polypeptide (encoded within an open reading frame of a CA sequence selected from any of the 95 polynucleotide sequences as mentioned in the specification, or its complement), an isolated antibody, (or its antigen binding fragment) that binds to the above polypeptide, a hybridoma that produces the above monoclonal antibody, a pharmaceutical composition comprising the above antibody and a pharmaceutical excipient, a kit for detecting cancer cells (comprising the antibody cited above, methods for diagnosing cancer or for detecting the presence or absence of cancer cells in an individual, a method for inhibiting growth of cancer cells in an individual, a method for delivering a therapeutic agent to cancer cells in an individual, an electronic library comprising the above polynucleotide or polypeptide for their fragments), methods of screening for anticancer activity or for a bioactive agent capable of modulating the activity of a CA protein (CAP), methods for detecting cancer associated with expression of a polypeptide in a test cell sample, a method for treating cancers and a method for inhibiting the expression of CA gene in a cell. The composition and methods are useful for detecting, diagnosing, preventing and treating cancers, especially lymphoma and

CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a human CAP cDNA sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 2215 BP; 420 A; 689 C; 665 G; 441 T; 0 U; 0 Other;

Query Match	4.9%;	Score 55.6;	DB 13;	Length 2215;
Best Local Similarity	49.3%;	Pred. No. 0.12;		
Matches 145;	Conservative 0;	Mismatches 149;	Indels 0;	Gaps 0;

Qy	325	GCGCGTGGCAGCGAGCCCGAGGATCACTGGCTTCTTCGACCGCAACCCCTGGACCTGGA	384
Db	83	GCCACCCCGCTGCGCCCGCCGCCCATGAGCGCGCGCGCTGCGCGCGGTCCT	142
Qy	385	GCCCGCGCTGCCACCCGCGCAGCGGGTGATCCTGCGCCACCCGAGGAGCAGCCTCTTTTC	444
Db	143	TCCGCTGCTGCTCGGGGCTTTCGCTGCTGGCGCGGAGTGGACGCGGATGCTCT	202
Qy	445	TGCGCGCTGGGGCCCGCCTCTCCATCCACAGTCTGGAGGCTCAGAGCCTGCGGTGCTGCA	504
Db	203	CCTGAGGCGCTGCTGTGCGGACGGACACCGGATGGCCACTCATCAGAAAGGACTGCTCGCT	262
Qy	505	GCCCTTCTGTACCCAGGACACGCGGATAGGCTTTTCAGGCGCAGGCCAGGAGGCT	564
Db	263	GCCATATGCTACGGATCCAAAGATGCAAGATGTCAGGAGCAGTGTGTCACAGCCA	322
Qy	565	GGACGTGCTGTGCGGCACCCCTCAGGCTGGTGGTGGTGGAGAACGAAGACCG	618
Db	323	GCTGAGGAGCTGCACTGTGCCACGGGATCAGCGCTGCGCCCAACGAGCAGGACCG	376

Search completed: May 30, 2005, 19:22:07
Job time : 978 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 15:53:13 ; Search time 8111 Seconds
(without alignments)
6762.582 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agccatggcagggccccgat.....agcagtgcgagcgagatcc 1132

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 segs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		Description	
Result No.	Score	Match	Length	DB	ID		
1	1132	100.0	1132	9	AF532984	AF532984 Homo sapi	
2	1132	100.0	1446	9	AB097667	AB097667 Homo sapi	
3	1116	98.6	1116	9	AF539796	AF539796 Homo sapi	
4	1116	98.6	1129	9	AF532983	AF532983 Homo sapi	
5	1116	98.6	1587	9	BC015917	BC015917 Homo sapi	
6	1107	97.8	1147	9	AF532985	AF532985 Homo sapi	
7	1100	97.2	1113	9	AY255768	AY255768 Homo sapi	
8	1091	96.4	1144	9	AY191359	AY191359 Homo sapi	
9	526	46.5	1260	10	BC019525	BC019525 Mus muscu	
10	519.6	45.9	1050	10	AF539797	AF539797 Mus muscu	
11	415.4	36.7	759	10	AB105906	AB105906 Cavia por	
12	410.4	36.3	412	6	CQ730806	CQ730806 Sequence	
13	378.8	33.5	720	10	AF399754	AF399754 Mus muscu	
C 14	344	30.4	80662	9	AC005606	AC005606 Homo sapi	
C 15	342.4	30.2	2922	6	AX834341	AX834341 Sequence	
C 16	342.4	30.2	2922	6	AK096738	AK096738 Homo sapi	
C 17	291.6	25.8	176290	2	AC148564	AC148564 Papio anu	
C 18	291.6	25.8	185895	2	AC150732	AC150732 Papio anu	
C 19	201	17.8	4948	9	AY212269	AY212269 Homo sapi	

20	159.8	14.1	884	5	BX933715	BX933715 Gallus ga
C 21	97.6	8.6	221847	2	AC127929	AC127929 Rattus no
C 22	97.6	8.6	231524	2	AC115181	AC115181 Rattus no
C 23	97.6	8.6	239810	2	AC111901	AC111901 Rattus no
C 24	97.6	8.6	265928	2	AC093937	AC093937 Rattus no
C 25	85.4	7.5	205691	2	AC087227	AC087227 Mus muscu
C 26	85.4	7.5	205737	2	AC021445	AC021445 Mus muscu
C 27	83	7.3	125020	9	AF429315	AF429315 Homo sapi
C 28	80	7.1	125020	9	AF429315	AF429315 Homo sapi
C 29	76.4	6.7	176704	2	CR628364	CR628364 Danio rer
C 30	70.8	6.3	1393	11	PM11H12G	AL684264 Penicilli
C 31	68.6	6.1	976	11	PM12D6G	AL684370 Penicilli
C 32	68.2	6.0	155337	2	AC116408	AL116408 Mus muscu
C 33	68	6.0	207420	2	AC078884	AC078884 Mus muscu
C 34	67	5.9	300695	2	AC079431	AC079431 Mus muscu
C 35	66.4	5.9	136889	2	AC135296	AC135296 Rattus no
C 36	65	5.7	95209	2	AP004323	AP004323 Oryza sat
C 37	64.8	5.7	885	11	PM7F8G	AL685189 Penicilli
C 38	64.8	5.7	85434	2	AC066610	AC066610 Homo sapi
C 39	64.6	5.7	136551	2	AC048354	AC048354 Homo sapi
C 40	64.6	5.7	136753	2	AC149968	AC149968 Strongylo
C 41	63.6	5.6	158053	2	CR293506	CR293506 Danio rer
C 42	61.8	5.5	991	11	PM12H12B	AL684455 Penicilli
C 43	61.6	5.4	1065	11	PM2B12B	AL684695 Penicilli
C 44	61.6	5.4	1094	11	PM7G11B	AL685196 Penicilli
C 45	61.4	5.4	976	11	PM12D6G	AL684370 Penicilli

ALIGNMENTS

RESULT 1	AF532984	1132 bp	mRNA	linear	PRI 02-FEB-2004
LOCUS	Homo sapiens regulatory protein NOX01-beta mRNA, complete cds;				
DEFINITION	alternatively spliced.				
ACCESSION	AF532984				
VERSION	AF532984.1	GI:22532412			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1132)				
AUTHORS	Cheng, G. and Lambeth, J.D.				
TITLE	NOX01, Regulation of Lipid Binding, Localization, and Activation of				
JOURNAL	Nox1 by the Phox Homology (PX) Domain				
PUBMED	J. Biol. Chem. 279 (6), 4737-4742 (2004)				
REFERENCE	2 (bases 1 to 1132)				
AUTHORS	Lambeth, J.D. and Cheng, G.				
TITLE	NOX01-beta: A homolog of p47phox				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 1132)				
AUTHORS	Lambeth, J.D. and Cheng, G.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JUL-2002) Pathology, Emory University Medical School,				
FEATURES	615 Michael St., Atlanta, GA 30322, USA				
source	Location/Qualifiers				
	1. .1132				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="16"				
	/map="16p13.3"				
	5. .1120				
	/note="similar to p47phox; alternatively spliced"				
	/codon_start=1				
	/product="regulatory protein NOX01-beta"				
	/protein_id="AA097926.1"				
	/db_xref="GI:22532413"				
	/translation="MAGRPYVSVOGAAVLVQIKRLQTFAFSVRWSDGDTFVRRSRL				
	FRQLKTKETFPVEAGLRDRVLPKLLDAPLLGRVGRGLRGLLLETYSRL				

CDS

1	1132				
2	1132				
3	1132				
4	1132				
5	1132				
6	1132				
7	1132				
8	1132				
9	1132				
10	1132				
11	1132				
12	1132				
13	1132				
14	1132				
15	1132				
16	1132				
17	1132				
18	1132				
19	1132				

LATARVARSPITTFAPQDLLEPALPPGSRVILPTPEOPLSRAAGRLSHSLEA
QSLRCLQPDCTQDTRDFPOAQOESLDVLLRHPSGWLVENEDROQTAFWPAPYLEEA
APQGRREGSILGSSGPFQASRAYESRADELSPAGARVRVLETSDRGHMLCRYGD
RAGLLPALLRPEGLGALUSGTGFRGGDDPAGEARGFPFESQATAPPVTPTRPSPGA
IQSRCCCTVTRALERRRQRPGRGCVDSVPHPTTEQ"
17. .370
/note="Region: PX domain"
485. .661
/note="Region: SH3 domain"
707. .874
/note="Region: SH3 domain"

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 100.0%; Score 1132; DB 9; Length 1132;
Best Local Similarity 100.0%; Pred. No. 7.3e-177; Indels 0; Gaps 0;
Matches 1132; Conservative 0; Mismatches 0;

QY 1 AGCCATGGCAGGCCCCCGATACCCAGTTCAGTGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
DB 1 AGCCATGGCAGGCCCCCGATACCCAGTTCAGTGTCAAGGGGCGAGCCCTGGTGCAGATCAA 60
QY 61 GAGGCTCAAAAGCTTTGCTTCTGTGGCTGTGTGAGGAGCGACACCTTCCTGTCGG 120
DB 61 GAGGCTCAAAAGCTTTGCTTCTGTGGCTGTGTGAGGAGCGACACCTTCCTGTCGG 120
QY 121 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAGACCTCAAGGAGACCTTCCTGTCGA 180
DB 121 CAGGAGTTGGGACGAATTCAGGAGCTCAAGAGACCTCAAGGAGACCTTCCTGTCGA 180
QY 181 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTTCGATGACCACTGTT 240
DB 181 GCGGGCTGTCTGCGAGATCTGACCGGTTCTCCAAAGCTTTCGATGACCACTGTT 240
QY 241 GGGAGCGTGGGGCGCAAGAGCGGCGCTGCGCGCTGTGAGCTGTGTGGAAACCTATTC 300
DB 241 GGGAGCGTGGGGCGCAAGAGCGGCGCTGCGCGCTGTGAGCTGTGTGGAAACCTATTC 300
QY 301 TCGGAGGCTGTGGCGACTGTGAGAGCGGTGTGACGGAGCCGACGATCACTGGCTTCT 360
DB 301 TCGGAGGCTGTGGCGACTGTGAGAGCGGTGTGACGGAGCCGACGATCACTGGCTTCT 360
QY 361 CGCACCGCAACCCCTGGACCTGGAGCCGCGCTGTCACCCGCGAGCCGGTGATCCTGCC 420
DB 361 CGCACCGCAACCCCTGGACCTGGAGCCGCGCTGTCACCCGCGAGCCGGTGATCCTGCC 420
QY 421 CACCCAGAGAGAGAGCTTTTCTGCGCTGTGCGGGCGGCTTCCATCCACAGTCTGGA 480
DB 421 CACCCAGAGAGAGAGCTTTTCTGCGCTGTGCGGGCGGCTTCCATCCACAGTCTGGA 480
QY 481 GGCTCAGAGCTTGCCTGCAGCCCTTCTGTATACCCAGAGACGCGGGATAGGCTTTT 540
DB 481 GGCTCAGAGCTTGCCTGCAGCCCTTCTGTATACCCAGAGACGCGGGATAGGCTTTT 540
QY 541 TCAGGCGCAGGCCCCAGGAGAGCTTGAGAGCTGTGCTGCGGGCACCCCTCAGGCTGGTGCT 600
DB 541 TCAGGCGCAGGCCCCAGGAGAGCTTGAGAGCTGTGCTGCTGCGGGCACCCCTCAGGCTGGTGCT 600
QY 601 GGTGAGAACCAAGAACCCGCGAGACCGCTGTGTTTCCAGCGCCCTTACCTGAGAGAGCGCG 660
DB 601 GGTGAGAACCAAGAACCCGCGAGACCGCTGTGTTTCCAGCGCCCTTACCTGAGAGAGCGCG 660
QY 661 CCGGGCCCAAGCCGGGAGGAGGCGCCCTTCCATAGGAGAGCGGTTCGCCAGTTCGTGTC 720
DB 661 CCGGGCCCAAGCCGGGAGGAGGCGCCCTTCCATAGGAGAGCGGTTCGCCAGTTCGTGTC 720
QY 721 TTCCCGCGCTACGAGAGCAGCCGCGCAGATGAGCTGTCCGTGCGCGGGGGCGCGCT 780
DB 721 TTCCCGCGCTACGAGAGCAGCCGCGCAGATGAGCTGTCCGTGCGCGGGGGCGCGCT 780
QY 781 GCGCGTGTGGAAACGTCAGACCGCGGCTGTGGCTATGCAAGGTACGCGGACCGGGCGGG 840
DB 781 GCGCGTGTGGAAACGTCAGACCGCGGCTGTGGCTATGCAAGGTACGCGGACCGGGCGGG 840

QY 841 CCTACTCCCCCGTGTGCTGCGCGGAAAGGCTGGCGCTCTCTCTGAGCGGACGGG 900
DB 841 CCTACTCCCCCGTGTGCTGCGCGGAAAGGCTGGCGCTCTCTCTGAGCGGACGGG 900
QY 901 GTTCCGTGGAGAGACGACACCGCGGGGTGAGCGCGGGCTTCCCTGAACCTTCCAGGC 960
DB 901 GTTCCGTGGAGAGACGACACCGCGGGGTGAGCGCGGGCTTCCCTGAACCTTCCAGGC 960
QY 961 CACGCGCCCTCCCCCGACCGGTGCCACCGACCTTCGCGGGGCGCATCCAGAGCGGTG 1020
DB 961 CACGCGCCCTCCCCCGACCGGTGCCACCGACCTTCGCGGGGCGCATCCAGAGCGGTG 1020
QY 1021 CTGCACCGTCAACGACGCGGCGCTGTGAGCGCGGCCACCGCGCGCGCCCTCCAGG 1080
DB 1021 CTGCACCGTCAACGACGCGGCGCTGTGAGCGCGGCCACCGCGCGCGCCCTCCAGG 1080
QY 1081 GTGCGTGGACTCTGTGCGGACCCACGACGAGCAGTGAGCGGAGGATCC 1132
DB 1081 GTGCGTGGACTCTGTGCGGACCCACGACGAGCAGTGAGCGGAGGATCC 1132

RESULT 2

AB097667

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

ORIGIN

Query Match 100.0%; Score 1132; DB 9; Length 1446;
Best Local Similarity 100.0%; Pred. No. 6.8e-177;

AB097667 1446 bp mRNA linear PRI 03-JUL-2003
Homo sapiens NOX01 mRNA for NADPH oxidase organizer 1, complete cds.
AB097667
AB097667.1 GI:30844229
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Takeya,R., Ueno,N., Kami,K., Taura,M., Kohjima,M., Izaki,T., Nunoi,H. and Sumimoto,H.
Novel human homologues of p47phox and p67phox participate in activation of superoxide-producing NADPH oxidases
J. Biol. Chem. 278 (27), 25234-25246 (2003)
12716264
12716910
2 (bases 1 to 1446)
Takeya,R., Taura,M. and Sumimoto,H.
Direct Submission
Submitted (05-DEC-2002) Ryo Takeya, Mediacal Institute of Bioregulation; 3-1-1, Maidashi, Fukuoka, Higashi-ku 812-8582, Japan
(E-mail:takeya@bioreg.kyushu-u.ac.jp. Tel:81-92-642-6811, Fax:81-92-642-6813)
Location/Qualifiers
1. .1446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .1446
/gene="NOX01"
181. .1296
/gene="NOX01"
/note="p4inox, a homologue of p47phox"
/codon_start=1
/product="NADPH oxidase organizer 1"
/protein_id="BAC76711.1"
/db_xref="GI:30844230"
translation="MAGPRPVSVQGAALVQIKELQTFAPSVRWSGSDTFFVRSSWDE
FRQKLTKETTFVEAGLRRSDRVLPALDLADPLLRGRTSRGLARQLLETYSRRL
LATARVARSPITTFAPQDLLEPALPPGSRVILPTPEOPLSRAAGRLSHSLEA
QSLRCLQPDCTQDTRDFPOAQOESLDVLLRHPSGWLVENEDROQTAFWPAPYLEEA
APQGRREGSILGSSGPFQASRAYESRADELSPAGARVRVLETSDRGHMLCRYGD
RAGLLPALLRPEGLGALUSGTGFRGGDDPAGEARGFPFESQATAPPVTPTRPSPGA
IQSRCCCTVTRALERRRQRPGRGCVDSVPHPTTEQ"

Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	AGCCATGGCAGGCCCCCGATACCCAGTTCACGTCGAAGGCGGAGCCCTCGTGAGATCAA	60						
DB	177	AGCCATGGCAGGCCCCCGATACCCAGTTCACGTCGAAGGCGGAGCCCTCGTGAGATCAA	236						
QY	61	GAGGCTCCAAACGTTTTCCTTCTCTGTGGCTGGTTCAGACGGGAGGACACCTTCGTGGG	120						
DB	237	GAGGCTCCAAACGTTTTCCTTCTCTGTGGCTGGTTCAGACGGGAGGACACCTTCGTGGG	296						
QY	121	CAGGAGTTCGGACGAATTCAGGCAGCTCAAGAAAGACCTCAAGGAGACCTTCCCGTGGG	180						
DB	297	CAGGAGTTCGGACGAATTCAGGCAGCTCAAGAAAGACCTCAAGGAGACCTTCCCGTGGG	356						
QY	181	GGCGGCTCTGTCGGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACCACTGT	240						
DB	357	GGCGGCTCTGTCGGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACCACTGT	416						
QY	241	GGACCGGTGGGCGGACAGCGCGGCTTGGCGGCTGAGCTGTGGAAACCTATTTC	300						
DB	417	GGACCGGTGGGCGGACAGCGCGGCTTGGCGGCTGAGCTGTGGAAACCTATTTC	476						
QY	301	TCGGAGGCTGCTGGGACCTGCAGAGCGCTGGCACGGAGCCGACCATCTGGCTTCT	360						
DB	477	TCGGAGGCTGCTGGGACCTGCAGAGCGCTGGCACGGAGCCGACCATCTGGCTTCT	536						
QY	361	CGCACCGCAACCCCTGGACCTGGAGCCCGGCTGCCACCGGAGCGGGTATCTCGC	420						
DB	537	CGCACCGCAACCCCTGGACCTGGAGCCCGGCTGCCACCGGAGCGGGTATCTCGC	596						
QY	421	CACCCAGAGGAGACCTTCTTCGCGTGGCGGCGGCTTCCATCCACAGTCTGA	480						
DB	597	CACCCAGAGGAGACCTTCTTCGCGTGGCGGCGGCTTCCATCCACAGTCTGA	656						
QY	481	GGCTCAGAGCTCGCTGCTCAGCGCTTCTGTACCCAGGACACGGGGATAGGCTTT	540						
DB	657	GGCTCAGAGCTCGCTGCTCAGCGCTTCTGTACCCAGGACACGGGGATAGGCTTT	716						
QY	541	TCAGGCGCAGGCGCCAGGAGAGCTGACCTGCTGCTGGCGCACCCCTCAGGCTGGGCT	600						
DB	717	TCAGGCGCAGGCGCCAGGAGAGCTGACCTGCTGCTGGCGCACCCCTCAGGCTGGGCT	776						
QY	601	GGTGAGAAACGAAGACCGGACACCGCTTGTTCAGGCGGCTTCTCAGGAGGCGGC	660						
DB	777	GGTGAGAAACGAAGACCGGACACCGCTTGTTCAGGCGGCTTCTCAGGAGGCGGC	836						
QY	661	CCGGGCGCAAGCGCGGAGGAGGCGGCTTCTCAGGAGGAGGCTTCTGCTG	720						
DB	837	CCGGGCGCAAGCGCGGAGGAGGCGGCTTCTCAGGAGGAGGCTTCTGCTG	896						
QY	721	TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGCTCGCGCGGCGGCGCGCT	780						
DB	897	TTCCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGCTCGCGCGGCGGCGCGCT	956						
QY	781	GGCGGTGTTGAAACGTCAGACCGCGCTGGTGGCTATGACGTAACCGCGCGG	840						
DB	957	GGCGGTGTTGAAACGTCAGACCGCGCTGGTGGCTATGACGTAACCGCGCGG	1016						
QY	841	CCTATCCTCCCGGCTGCTGCTGGCGCGGAGAGGCTGGGGCTCTCTCAGCGGAGCGG	900						
DB	1017	CCTATCCTCCCGGCTGCTGCTGGCGCGGAGAGGCTGGGGCTCTCTCAGCGGAGCGG	1076						
QY	901	GTTTCGTGAGAGAGACACCGCGGCTGAGGCGCGGGCTTCCCTGAACCTTCCAGGC	960						
DB	1077	GTTTCGTGAGAGAGACACCGCGGCTGAGGCGCGGGCTTCCCTGAACCTTCCAGGC	1136						
QY	961	CACCGGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1020						
DB	1137	CACCGGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1196						
QY	1021	CTGACCGTCAACGAGGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1080						
DB	1197	CTGACCGTCAACGAGGCGGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1256						

QY	1081	GTGCGTGGACTCTGTGCGCACCCACGACGAGCAGTGCAGCGGAGATCC	1132						
Db	1257	GTGCGTGGACTCTGTGCGCACCCACGACGAGCAGTGCAGCGGAGATCC	1308						
RESULT 3									
AF539796	Homo sapiens NADPH oxidase organizer 1 (NOXO1) mRNA, complete cds.								
LOCUS	AF539796	1116 bp	mRNA	linear	PRI 03-FEB-2003				
DEFINITION	Homo sapiens NADPH oxidase organizer 1 (NOXO1) mRNA, complete cds.								
ACCESSION	AF539796								
VERSION	AF539796.1 GI:25573147								
KEYWORDS									
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
AUTHORS	Banfi,B., Clark,R.A., Steger,K. and Krause,K.-H.								
TITLE	Two Novel Proteins Activate Superoxide Generation by the NADPH Oxidase NOX1								
JOURNAL	J. Biol. Chem. 278 (6), 3510-3513 (2003)								
PUBMED	12473664								
REFERENCE	2 (bases 1 to 1116)								
AUTHORS	Banfi,B., Clark,R.A., Steger,K. and Krause,K.-H.								
TITLE	Direct Submission								
JOURNAL	Submitted (21-AUG-2002) Dept. of Geriatrics, University Hospitals of Geneva, Chemin du Petit-Bel-Air 2, Geneva CH-1225, Switzerland								
FEATURES	Location/Qualifiers								
source	1..1116								
	/organism="Homo sapiens"								
	/mol_type="mRNA"								
	/db_xref="taxon:9606"								
gene	1..1116								
CDS	1..1116								
	/gene="NOXO1"								
	/note="p41NOX"								
	/codon_start=1								
	/product="NADPH oxidase organizer 1"								
	/protein_id="AA75141.1"								
	/db_xref="GI:25573148"								
	/translation="MAGPRYPVSVQGAALVQIKLQTFAPSVRSDGTFVRRSRWDE FRLKXKTKETFPVAGLLRSDVLPKLLDAPLGRVTRSGRLARLQLLETYSRRL LQATRVARSPTTGFPAQPLDLEPALPGSRVILPTPEQPLSRAAGRLSTHLEA QSLCRQPGCTQDTRDRFPQAOQSLDVLRLRHPSPWLVENEDRQTAFWPAPYLEA APQREGGSLGSSGPOFCASRAYESSRADELSPAGARVRVLETSDRGWLCRYGD RAGLLPALLRPEGLGALLSGTFRGGDDPAGEARGFPEPSQATAPPVTPTRPSGA IQSRCTVTRALERRRRQRRPGRCVDSVPHPPTTEQ"								
ORIGIN									
Query Match 98.6%; Score 1116; DB 9; Length 1116;									
Best Local Similarity 100.0%; Pred. No. 3.2e-174;									
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	5	ATGGCAGGCCCCCGATACCAGTTTCAGTGCAGGCGGAGCCCTGGTGAGATCAAGAGG	64						
Db	1	ATGGCAGGCCCCCGATACCAGTTTCAGTGCAGGCGGAGCCCTGGTGAGATCAAGAGG	60						
QY	65	CTCCAAACGTTTTCCTTCTCTGTGGCTGGTTCAGACGGCAGCACCTTCGTGCGCAGG	124						
Db	61	CTCCAAACGTTTTCCTTCTCTGTGGCTGGTTCAGACGGCAGCACCTTCGTGCGCAGG	120						
QY	125	AGTTGGGAGCAATTACGGAGCTCAAGAGACCCCTCAAGAGACCTTCCCGGTGGAGGGC	184						
Db	121	AGTTGGGAGCAATTACGGAGCTCAAGAGACCCCTCAAGAGACCTTCCCGGTGGAGGGC	180						
QY	185	GGCTGCTCGGAGATCTCACCGCTTCTCCCAAAGCTTCTCGATGCACCACTGTGGGA	244						
Db	181	GGCTGCTCGGAGATCTCACCGCTTCTCCCAAAGCTTCTCGATGCACCACTGTGGGA	240						
QY	245	CGCGTGGGCGCACGAGCGCGGCTTGGCGGCTTGTGAAACCTATTCTCGG	304						

Db 241 CGCGTGGGCGCACGAGCCGCGGCTGTGGCGGCTGTGAGCTGTGTGAAACCTATTCTCGG 300
Qy 305 AGGCTGCTGGGACTGTCAGAGGCGTGTGACGGAGCCGACGATCACTGGCTTCTTCGCA 364
Db 301 AGGCTGCTGGGACTGTCAGAGGCGTGTGACGGAGCCGACGATCACTGGCTTCTTCGCA 360
Qy 365 CGCAAAACCCCTGGACCTTGAGGCGCGCTGTCCACCCGCGAGCGGGGTGATCTGCCACCC 424
Db 361 CGCAAAACCCCTGGACCTTGAGGCGCGCTGTCCACCCGCGAGCGGGGTGATCTGCCACCC 420
Qy 425 CCAGAGGACGACCTCTTTCTCGCGCTGTGGGCGGCTCTCCATCACAGTCTGAGGCT 484
Db 421 CCAGAGGACGACCTCTTTCTCGCGCTGTGGGCGGCTCTCCATCACAGTCTGAGGCT 480
Qy 485 CAGAGCTGCGTGTGAGGCGCTCTGATACCGAGGACAGCGGGATAGGCTTTTCAG 544
Db 481 CAGAGCTGCGTGTGAGGCGCTCTGATACCGAGGACAGCGGGATAGGCTTTTCAG 540
Qy 545 GCGCAGGCGCCAGGAGGCTTGACGCTGTCTGTGGGCGACCCCTCAGGCTGGTGGCTGGTG 604
Db 541 GCGCAGGCGCCAGGAGGCTTGACGCTGTCTGTGGGCGACCCCTCAGGCTGGTGGCTGGTG 600
Qy 605 GAGAACGAAGACCCGCGAGACCGCTGTGTTTCAGCGCCCTACCTGAGGAGCGCGCCCG 664
Db 601 GAGAACGAAGACCCGCGAGACCGCTGTGTTTCAGCGCCCTACCTGAGGAGCGCGCCCG 660
Qy 665 GCGCAAGCGCGGAGGAGGCGGCTCTCTAGGAGCAGCGGTCGCCAGTCTCTGTCTTCC 724
Db 661 GCGCAAGCGCGGAGGAGGCGGCTCTCTAGGAGCAGCGGTCGCCAGTCTCTGTCTTCC 720
Qy 725 CGCGCTACGAGAGCAGCGCGCAGATGATGCTGCTGCGGCGCGCGCGGCTGGCG 784
Db 721 CGCGCTACGAGAGCAGCGCGCAGATGATGCTGCTGCGGCGCGCGCGGCTGGCG 780
Qy 785 GTGTTGAAAACGTCAGACCGCGGCTGTGGCTATGACAGGTAGCGACCGCGCGGCGCTA 844
Db 781 GTGTTGAAAACGTCAGACCGCGGCTGTGGCTATGACAGGTAGCGACCGCGCGGCGCTA 840
Qy 845 CTCGCCGCGTGTCTGTCGCGCGGAGGCGTGGGCGCTCTCTGAGCGGAGCGGGTTC 904
Db 841 CTCGCCGCGTGTCTGTCGCGCGGAGGCGTGGGCGCTCTCTGAGCGGAGCGGGTTC 900
Qy 905 CTGAGAGAGACGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTTCCAGGCGCAC 964
Db 901 CTGAGAGAGACGACCCGCGGCTGAGGCGCGGCTTCCCTGAACCTTCCAGGCGCAC 960
Qy 965 GCCCTCCCGCCACCGTGTCCACCGACCTTGTGCGGCGGCGCATCCAGAGCGCTGTGTC 1024
Db 961 GCCCTCCCGCCACCGTGTCCACCGACCTTGTGCGGCGGCGCATCCAGAGCGCTGTGTC 1020
Qy 1025 ACCGTACAGCAGGCGGCTGTGAGCGGCGGCGCATCCAGGCGCGGCTGTGAGGCTGC 1084
Db 1021 ACCGTACAGCAGGCGGCTGTGAGCGGCGGCGCATCCAGGCGCGGCTGTGAGGCTGC 1080
Qy 1085 GTGAGCTCTGTGCGGCGACCCCGACGAGCAGTGA 1120
Db 1081 GTGAGCTCTGTGCGGCGACCCCGACGAGCAGTGA 1116

RESULT 4
AF532983
LOCUS AF532983 1129 bp mRNA linear PRI 03-JAN-2003
DEFINITION Homo sapiens regulatory protein NOX01-alpha mRNA, complete cds;
alternatively spliced.
ACCESSION AF532983
VERSION AF532983.1 GI:22532410
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1129)
AUTHORS Lambeth,J.D. and Cheng,G.

NOX01-alpha: A homolog of p47phox
Unpublished
2 (bases 1 to 1129)
Lambeth,J.D. and Cheng,G.
Direct Submission
Submitted (29-JUL-2002) Pathology, Emory University Medical School,
615 Michael St., Atlanta, GA 30322, USA
FEATURES
source
1..1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.3"
5..1117
/note="similar to p47phox; alternatively spliced"
/codon_start=1
/product="regulatory protein NOX01-alpha"
/protein_id="AAM97925.1"
/db_xref="GI:22532411"
/translation="MAGPRYPVSVQGAALVQIKELQTFAESVRMSDGSDFTVRSWDE
FRQKTKETFPVEAGLRSDRVLPKLDAPLGRVGRSLRGLARLQLLTYSRRLL
ATAERVARSPTITGFAPQPDLEPALPPGSRVILPTPEQPLSRAARGLPYSRLLEAQ
SLRCLQFCTQDTRDPFQAQESLDVLRHPSGWMVENEDQTAFFPAPYLEEA
PGQREGPSILSGSQPOFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDR
AGLLPAVLLRPEGLALLSGTGRGDDPAGARGFPEPSQATAPPTVTPTRPSGAI
QSRCTVTRALERPRRQGRGCVDSVPHPTTEQ"
17..367
/note="Region: PX domain"
482..658
/note="Region: SH3 domain"
704..880
/note="Region: SH3 domain"
ORIGIN
Query Match 98.6%; Score 1116; DB 9; Length 1129;
Best Local Similarity 99.7%; Pred. No. 3.1e-174;
Matches 1129; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 1 AGCCATGCGAGGCCCGATACCCAGCTTTCAGTCAAGGGGCGAGCCCTGTGGCAGATCAA 60
Db 1 AGCCATGCGAGGCCCGGATACCCAGCTTTCAGTCAAGGGGCGAGCCCTGTGGCAGATCAA 60
Qy 61 GAGGCTCCAAACGTTTGCCTTCTCTGTCGCTGTGTCAGCGCAGCGCAGCACCTTCTGTCG 120
Db 61 GAGGCTCCAAACGTTTGCCTTCTCTGTCGCTGTGTCAGCGCAGCGCAGCACCTTCTGTCG 120
Qy 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAGACCTTCAAGGAGACCTTCCCGGTGGA 180
Db 121 CAGGAGTTGGGACGAATTCAGGCGAGCTCAAGAAGACCTTCAAGGAGACCTTCCCGGTGGA 177
Qy 181 GCGGGGCTGTGCGGAGACTGACCGGTTCTCCAAAGCTTTCGATGACCACTGTT 240
Db 178 GCGGGGCTGTGCGGAGACTGACCGGTTCTCCAAAGCTTTCGATGACCACTGTT 237
Qy 241 GGGACGCTGGGGCGCAGCGCGCTGCGCGCTGCGAGCTGTGGAAACTATTC 300
Db 238 GGGACGCTGGGGCGCAGCGCGCTGCGCGCTGCGAGCTGTGGAAACTATTC 297
Qy 301 TCGGAGGCTGTGCGGAGCTGCGAGCGCTGGCAGCGGACCGGACCGGACCTTCTGCTTCTT 360
Db 298 TCGGAGGCTGTGCGGAGCTGCGAGCGCTGGCAGCGGACCGGACCGGACCTTCTGCTTCTT 357
Qy 361 CGCACCAGAACCCCTGGACCTGGAGCGCGCTGCCACCCCGGAGCGGGGTATCCTGCC 420
Db 358 CGCACCAGAACCCCTGGACCTGGAGCGCGCTGCCACCCCGGAGCGGGGTATCCTGCC 417
Qy 421 CACCCAGAGGAGCAGCTCTTTCTGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGA 480
Db 418 CACCCAGAGGAGCAGCTCTTTCTGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGA 477
Qy 481 GGCTCAGAGCCTGCGCTGCGCTGCGAGCCTTCTGTACCCAGGACACGCGGGATAGCCCTTT 540

Db 257 AGCATGGCAGGCCCCCGATACCAGTTTTCAGTGCAGAGGGCAGCCCTGGTGCAGATCAA 316

Qy 61 GAGGCTCCAAAGCTTTGCTTCTGTGGCTGGTTCAGACGGCAGCGACACCTTGTGCG 120

Db 317 GAGGCTCCAAAGCTTTGCTTCTGTGCTGGTTCAGACGGCAGCGACACCTTGTGCG 376

Qy 121 CAGGAGTTGGGACGAATTCAGGCAGCTCAAGAAGACCTCAAGGAGACTTCCCGGTGGA 180

Db 377 CAGGAGTTGGGACGAATTCAGGCAGCTC---AAGACCTCAAGGAGACTTCCCGGTGGA 433

Qy 181 GCGGGCTGCTGCGGAGACTGACCGGTTCTCCAAAGCTTTCGATGACACACTGTT 240

Db 434 GCGGGCTGCTGCGGAGACTGACCGGTTCTCCAAAGCTTTCGATGACACACTGTT 493

Qy 241 GCGAGCTGGGGCCACGAGCGCGGCTGGCGGCTGCGGCTGCGAGCTTGTGGAACTATTC 300

Db 494 GCGAGCTGGGGCCACGAGCGCGGCTGCGGCTGCGGCTGCGAGCTTGTGGAACTATTC 553

Qy 301 TCGGAGGCTGCTGGCGACTGACAGAGCGGTGGACGGAGCCGACGATCACTGGCTTCTT 360

Db 554 TCGGAGGCTGCTGGCGACTGACAGAGCGGTGGACGGAGCCGACGATCACTGGCTTCTT 613

Qy 361 CGCACCGCAACCTTGGACCTGGAGCCCGGCTGCCACCCGCGCAGCCGCGGTGATCCTGCC 420

Db 614 CGCACCGCAACCTTGGACCTGGAGCCCGGCTGCCACCCGCGCAGCCGCGGTGATCCTGCC 673

Qy 421 CACCCACAGAGAGAGCTTCTTCTGCGCTGCGGGCCGCTTCTCCATCCACAGTCTGGA 480

Db 674 CACCCACAGAGAGAGCTTCTTCTGCGGCTGCGGGCCGCTTCTCCATCCACAGTCTGGA 733

Qy 481 GGCTCAGAGCTGCGCTGCTGCGCTTCTGTATCCAGAGACACGCGGGATAGSCCTTT 540

Db 734 GGCTCAGAGCTGCGCTGCTGCGCTTCTGTATCCAGAGACACGCGGGATAGSCCTTT 793

Qy 541 TCAGCGCAGGCCACAGGAGAGCTTGGAGCTGCTGCTGGGACACCCCTCAGGCTGTGCT 600

Db 794 TCAGCGCAGGCCACAGGAGAGCTTGGAGCTGCTGCTGGGACACCCCTCAGGCTGTGCT 853

Qy 601 GGTGAGAAACGAGACCGCGAGACCGCTGTTTTCAGCGGCTTACCTTGGAGAGGCGGC 660

Db 854 GGTGAGAAACGAGACCGCGAGACCGCTGTTTTCAGCGGCTTACCTTGGAGAGGCGGC 913

Qy 661 CCCGGCCAAAGCCCGGAGGAGGCGCTTCTAGGAGACAGCGTCCCACTTCTGTGC 720

Db 914 CCCGGCCAAAGCCCGGAGGAGGCGCTTCTAGGAGACAGCGTCCCACTTCTGTGC 973

Qy 721 TTCCGCGCCTACGAGAGACCGCGCAGATGAGTGTCCGTGCCCGGGGCGCGGT 780

Db 974 TTCCGCGCCTACGAGAGACCGCGCAGATGAGTGTCCGTGCCCGGGGCGCGGT 1033

Qy 781 GCGGCTGTTGGAACGTGACACCGCGGCTGTGGCTATGACAGTACGGCGACCGGGCGGG 840

Db 1034 GCGGCTGTTGGAACGTGACACCGCGGCTGTGGCTATGACAGTACGGCGACCGGGCGGG 1093

Qy 841 CCTACTCCCGCGGTGCTGCTGCGCCGGAAGGGCTGGCGCTCTCTGAGCGGACCGGG 900

Db 1094 CCTACTCCCGCGGTGCTGCTGCGCCGGAAGGGCTGGCGCTCTCTGAGCGGACCGGG 1153

Qy 901 GTTTCGTGGAGAGACGACCCCGGGGTGAGGCCCGGGGCTTCCCTGAACTCCACGCG 960

Db 1154 GTTTCGTGGAGAGACGACCCCGGGGTGAGGCCCGGGGCTTCCCTGAACTCCACGCG 1213

Qy 961 CACGCGCTTCCCTCCACCGTGCACCCGACCTTTCGCGGGCGCATCCAGAGCGGTG 1020

Db 1214 CACGCGCTTCCCTCCACCGTGCACCCGACCTTTCGCGGGCGCATCCAGAGCGGTG 1273

Qy 1021 CTGACCTTCACGACGAGGCGCTGAGCGGCGCCACGCGGCCAGGGCCGCTTCGAGG 1080

Db 1274 CTGACCTTCACGACGAGGCGCTGAGCGGCGCCACGCGGCCAGGGCCGCTTCGAGG 1333

Qy 1081 GTGCGTGGACTGTGTGCGGACACCCACGACGAGCAGTGTGAGCGCGAGGATCC 1132

Db 1334 GTGCGTGGACTGTGTGCGGACACCCACGACGAGCAGTGTGAGCGCGAGGATCC 1385

RESULT 6

AF532985

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 1147

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="16"

/map="16p13.3"

5. 1135

/note="similar to p47phox; alternatively spliced"

/codon_start=1

/product="regulatory protein NOXO1-gamma"

/protein_id="AA097927.1"

/db_xref="GI:22532415"

/translation="MAGPRYPVSVQGAALVOIKRLQTFAFSVRWSDGSDTFVRSWDE

PROLKTLLKTFPVEAGLRSDRLPKLLGOASLDAPLLGRVGRTRGLARLOLLET

YSRLLATAERVARSPITITGFPAQPLDLSPALPPGSRVILPTPEEPLSAAAGLSI

HSLSAQRLCQLPFTQDRPFAQAQSLDLVLRHPGSMWLVENEDRQTANFPAP

YLIEAAPGQREGFSLGSGPOFCASRAYESRDELVSVPAGARVLRVLETSDRGWLL

CTEADAGLLPALLRPEGLGALLSGTFRGGDDPAGEARGFPEPSQATAPPTVPTTR

PSPAIQSRCTVTTRALERPRRQGRPRGCVDSVPHTTREQ"

misc_feature

17. 385

/note="Region: PX domain"

misc_feature

500. 676

/note="Region: SH3 domain"

misc_feature

722. 889

/note="Region: SH3 domain"

ORIGIN

Query Match 97.8%; Score 1107; DB 9; Length 1147;

Best Local Similarity 98.7%; Pred. No. 9.5e-173;

Matches 1132; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Qy 1 AGCATGGCAGGCCCCCGATACCAGTTTCAGTGCAGAGGGCAGCCCTGGTGCAGATCAA 60

Db 1 AGCATGGCAGGCCCCCGATACCAGTTTCAGTGCAGAGGGCAGCCCTGGTGCAGATCAA 60

Qy 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGCGTGGTTCAGACGGCAGCGACACCTTCGTGCG 120

Db 61 GAGGCTCCAAAGCTTTGCTTCTCTGTGCGTGGTTCAGACGGCAGCGACACCTTCGTGCG 120

Qy 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Db 121 CAGGAGTTGGAGCGAATTCAGGCGAGCTCAAGAAGACCTTCAAGAGACCTTCCCGGTGGA 180

Qy 181 GCGCGGCTGCTGCGGAGACTGACCGGTTCTCCAAAGCTTCTC----- 226

Db 181 GCGCGGCTGCTGCGGAGACTGACCGGTTCTCCAAAGCTTCTC----- 240

Qy 227 -GATGCACCACTGTTGGGACGCGTGGGGCGCACGAGCGCGGCTTGGCGCGCTGCACT 285

|||||

```
Db 241 GGATGACCACTGTTGGAGCGCTGGGGCGCACAGAGCCGGCGCTCGCGCGCTGAGCT 300
Qy 286 GTTGGAAACCTATTCTCGAGGCTGTGGCGACTGCGAGCGCGTGGCAGCGAGCCCGAC 345
Db 301 GTTGGAAACCTATTCTCGAGGCTGTGGCGACTGCGAGCGCGTGGCAGCGAGCCCGAC 360
Qy 346 GATCACTGGCTTCTCGACCGCAACCCCTGAGACTGAGAGCCCGCGCTGCACCCGGCAG 405
Db 361 GATCACTGGCTTCTCGACCGCAACCCCTGAGACTGAGAGCCCGCGCTGCACCCGGCAG 420
Qy 406 CCGGGTGATCTGCCCCACCCAGAGAGCAGCTCTTTCTCGCGCTGCGGGCCGCTCTC 465
Db 421 CCGGGTGATCTGCCCCACCCAGAGAGCAGCTCTTTCTCGCGCTGCGGGCCGCTCTC 480
Qy 466 CATCCACAGTCTGAGGCTCAGAGCTCGGCTCGCTCGCAGAGCCCTTCTTACCCAGAGAC 525
Db 481 CATCCACAGTCTGAGGCTCAGAGCTCGGCTCGCTCGCAGAGCCCTTCTTACCCAGAGAC 540
Qy 526 GCGGGATAGCCCTTTTACGCGCAGGCCCGCAGAGAGCTTGAGCTGCTGCTGGCGCACCC 585
Db 541 GCGGGATAGCCCTTTTACGCGCAGGCCCGCAGAGAGCTTGAGCTGCTGCTGGCGCACCC 600
Qy 586 CTCAGGCTGCTGCTGCTGAGAACGAGACCGGCAGAGCGCTGCTGCTTCCAGCGCCCTA 645
Db 601 CTCAGGCTGCTGCTGCTGAGAACGAGACCGGCAGAGCGCTGCTGCTTCCAGCGCCCTA 660
Qy 646 CTGAGAGAGGCGGCCCGCGCCCAAGCGCGGAGGAGGCGCCCTCCCTAGGAGAGAGCGG 705
Db 661 CTGAGAGAGGCGGCCCGCGCCCAAGCGCGGAGGAGGCGCCCTCCCTAGGAGAGAGCGG 720
Qy 706 TCCCCAGTCTGCTTCCCGCCCTACGAGAGCAGCGCGCAGATGAGCTGCTCGTGC 765
Db 721 TCCCCAGTCTGCTTCCCGCCCTACGAGAGCAGCGCGCAGATGAGCTGCTCGTGC 780
Qy 766 CCGGGGGCGCGTGGCGGTGTTGAAACGTCAGACCGCGGCTGCTGCTATGAGGTA 825
Db 781 CCGGGGGCGCGTGGCGGTGTTGAAACGTCAGACCGCGGCTGCTGCTATGAGGTA 840
Qy 826 CCGCGACCGGGCGGGCTACTCTCCCGCGGTGCTGCTCGCGCGCGAAGGCTGGCGCTCT 885
Db 841 CCGCGACCGGGCGGGCTACTCTCCCGCGGTGCTGCTCGCGCGCGAAGGCTGGCGCTCT 900
Qy 886 CTGAGCGGAGCGGGTTCCTGGAGAGACGACCGCGGGTGAGCGCGGGCTTCCC 945
Db 901 CTTAGCGGAGCGGGTTCCTGGAGAGACGACCGCGGGTGAGCGCGGGCTTCCC 960
Qy 946 TGAACCTCCAGGCCACCGCCCTCCCGCCCGTGGCCACCGACCTTCGCGCGGGCGC 1005
Db 961 TGAACCTCCAGGCCACCGCCCTCCCGCCCGTGGCCACCGACCTTCGCGCGGGCGC 1020
Qy 1006 CATCCAGAGCGCTGTGCACCGTCACAGCAGGGGCTTGGAGCGCGCCACCGGCCCA 1065
Db 1021 CATCCAGAGCGCTGTGCACCGTCACAGCAGGGGCTTGGAGCGCGCCACCGGCCCA 1080
Qy 1066 GGGCGCCCTCGAGGGTGGTGAATCTGTGCGCGCACCCACAGCAGGAGCAGTGAAGCGG 1125
Db 1081 GGGCGCCCTCGAGGGTGGTGAATCTGTGCGCGCACCCACAGCAGGAGCAGTGAAGCGG 1140
Qy 1126 AGGATCC 1132
Db 1141 AGGATCC 1147
```

```
RESULT 7
LOCUS AY255768 1113 bp mRNA linear PRI 27-MAY-2003
DEFINITION Homo sapiens NADPH oxidase regulatory protein (NOXO1) mRNA,
complete cds.
ACCESSION AY255768
VERSION AY255768.1 GI:30102429
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Geiszt M., Lekstrom, K., Witta, J. and Leto, T.L.
TITLE Proteins Homologous to p47phox and p67phox Support Superoxide
PUBMED Production by NAD(P)H Oxidase 1 in Colon Epithelial Cells
J. Biol. Chem. 278 (22), 20006-20012 (2003)
12657628
REFERENCE
AUTHORS Geiszt M., Lekstrom, K. and Leto, T.L.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2003) Laboratory of Host Defenses, National
Institute of Allergy and Infectious Diseases, National Institutes
of Health, Building 10, Room 11 N 106, Bethesda, MD 20892, USA
FEATURES
source
1. .1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="IMAGE:4661469"
/tissue_type="colon; adenomcarcinoma"
1. .1113
/gene="NOXO1"
1. .1113
/gene="NOXO1"
/note="p4inox; NOXO1; similar to p47phox (NCF1); p4inox;
supports activity of NADPH oxidase 1 (Nox1)"
/codon_start=1
/product="NADPH oxidase regulatory protein"
/protein_id="AAP13479.1"
/db_xref="GI:30102430"
/translation="MAGPRYPVSVVOGAALVQIKRLQTPAFSVRWSDGSDTFVRRSWDE
PROKLTLETPEVEAGLRSDRLPKLLDAPLLGRVGRTSRLARLQLLETYSRLL
ATAERVASPIITGFPAPQPDLDLPALPGSRVILPTPEEQPLSRAAGRLSHSLEAQ
SLRCLQPTQDTRDPFQAQESLDVLLRHPSWGLVLENEDROTAMPFPALYLEEAQ
PQGRGEPFSLGSGQFCASRAYESRADELSVPAGARVRLVLETSDRGMWLCRYGDR
AGLPAVLLRPEGLALSLGTGFGDDPAGARGFPPEPSQATAPPPTVTPRSPGAI
QSRCCVTTRALLERRPRQRGRGCVDSVPHPTTEQ"
```

ORIGIN

```
Query Match 97.2%; Score 1100; DB 9; Length 1113;
Best Local Similarity 99.7%; Pred. No. 1.4e-171;
Matches 1113; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 5 ATGGCAGGCCCCCGATACCCAGTTTCAGTGCAAGGGGAGCCCTCGTCAGATCAAGAGG 64
Db 1 ATGGCAGGCCCCCGATACCCAGTTTCAGTGCAAGGGGAGCCCTCGTCAGATCAAGAGG 60
Qy 65 CTCCAAACGTTTGGCTTCTCTGTGCGCTGTCAGACGCGCAGCGACACCTTCGTGCGCAGG 124
Db 61 CTCCAAACGTTTGGCTTCTCTGTGCGCTGTCAGACGCGCAGCGACACCTTCGTGCGCAGG 120
Qy 125 AGTTGGAGCAATTCAGGAGCTCAAGAAGACCTTCAGGAGACCTTCCTCGGTGGAGGGG 184
Db 121 AGTTGGAGCAATTCAGGAGCTC---AAGACCTCAAGGAGACCTTCCTCGGTGGAGGGG 177
Qy 185 GGCCTGCTGCGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACACCTGTTTGGGA 244
Db 178 GGCCTGCTGCGAGATCTGACCGCTTCTCCAAAGCTTCTCGATGCACACCTGTTTGGGA 237
Qy 245 CGCGTGGGCGCACGAGCGCGCTTGGCGCGCTGCGACCTGTGGAACCTATTCTCGG 304
Db 238 CGCGTGGGCGCACGAGCGCGCTTGGCGCGCTGCGACCTGTGGAACCTATTCTCGG 297
Qy 305 AGGCTGTGGCGACTGCGAGAGCGCTGGCAGCGGAGCCGACGATCACTGGCTTCTTCGCA 364
Db 298 AGGCTGTGGCGACTGCGAGAGCGCTGGCAGCGGAGCCGACGATCACTGGCTTCTTCGCA 357
Qy 365 CCGCAACCCCTGGAAGCTGGAGCCCGCTGCGACCCGCGAGCGGGTGATCTCTGCCACC 424
Db 358 CCGCAACCCCTGGAAGCTGGAGCCCGCTGCGACCCGCGAGCGGGTGATCTCTGCCACC 417
```


QY 425 CCAGAGGAGCAGCTCTTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGAGGCT 484
| | | | |
Db 418 CCAGAGGAGCAGCTCTTTCTCGCGCTGCGGCGCGCTCTCCATCCACAGTCTGGAGGCT 477
| | | | |
QY 485 CAGAGCCTGCGCTGCTGAGCCCTTCTGTATCCCAAGGACACGCGGGATAGGCTTTTTCAG 544
| | | | |
Db 478 CAGAGCCTGCGCTGCTGAGCCCTTCTGTATCCCAAGGACACGCGGGATAGGCTTTTTCAG 537
| | | | |
QY 545 GCGCAGGCGCAGGAGAGCCTGAGCTGCTGTGCGGACACCCCTCAGGCTGTGTGCTGTG 604
| | | | |
Db 538 GCGCAGGCGCAGGAGAGCCTGAGCTGCTGTGCGGACACCCCTCAGGCTGTGTGCTGTG 597
| | | | |
QY 605 GAGAACGAGACCGCAGACCGCCTGTTTTCAGCGCCCTACTCTGAGGAGCGCGCCCG 664
| | | | |
Db 598 GAGAACGAGACCGCAGACCGCCTGTTTTCAGCGCCCTACTCTGAGGAGCGCGCCCG 657
| | | | |
QY 665 GSCCAAGCGCGGAGGAGCGCCCTCCCTAGGAGCAGCGGTCCTCAGTCTGTGCTTCC 724
| | | | |
Db 658 GSCCAAGCGCGGAGGAGCGCCCTCCCTAGGAGCAGCGGTCCTCAGTCTGTGCTTCC 717
| | | | |
QY 725 CCGCGCTAGAGAGCAGCGCGCAGATGAGCTGTCTGCTGCGCGGCGCGCGCTGCGC 784
| | | | |
Db 718 CCGCGCTAGAGAGCAGCGCGCAGATGAGCTGTCTGCTGCGCGGCGCGCGCTGCGC 777
| | | | |
QY 785 GTGTTGGAAACGTACAGACCGCGCTGTGGCTATGACAGTACGCGACCGGCGCGCCTA 844
| | | | |
Db 778 GTGTTGGAAACGTACAGACCGCGCTGTGGCTATGACAGTACGCGACCGGCGCGCCTA 837
| | | | |
QY 845 CTCCCGCGGCTGCTGCTGCGCGCGAAGGCTGCGCTCTCTCTGAGCGGACGCGGTTTC 904
| | | | |
Db 838 CTCCCGCGGCTGCTGCTGCGCGCGAAGGCTGCGCTCTCTCTGAGCGGACGCGGTTTC 897
| | | | |
QY 905 CBTGAGAGAGCAGACCCGCGGCTGAGCGCGCGCTTCCCTGAACCTCCAGGCGCAC 964
| | | | |
Db 898 CBTGAGAGAGCAGACCCGCGGCTGAGCGCGCGCTTCCCTGAACCTCCAGGCGCAC 957
| | | | |
QY 965 GCCCTTCCCGCCACCGTGCACCGACCTTCCCGCGGCGCATCCAGAGCGCTGCTGC 1024
| | | | |
Db 958 GCCCTTCCCGCCACCGTGCACCGACCTTCCCGCGGCGCATCCAGAGCGCTGCTGC 1017
| | | | |
QY 1025 ACCGTACACGAGGCGCTTGAGCGGCGCCCAAGCGCGCGCTTCCCTGAAGCTCCAGGCGTGC 1084
| | | | |
Db 1018 ACCGTACACGAGGCGCTTGAGCGGCGCCCAAGCGCGCGCTTCCCTGAAGCTCCAGGCGTGC 1077
| | | | |
QY 1085 GTGGACTCTGTGCGCGACCCACCGACGAGCAGTGA 1120
| | | | |
Db 1078 GTGGACTCTGTGCGCGACCCACCGACGAGCAGTGA 1113
| | | | |

RESULT 8
AY191359 1144 bp mRNA linear PRI 15-FEB-2003
LOCUS
DEFINITION Homo sapiens regulatory protein NOX01-delta (NOX01) mRNA, complete cds; alternatively spliced.
ACCESSION AY191359
VERSION AY191359.1 GI:28394745
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1144)
AUTHORS Cheng, G. and Lambeth, J.D.
TITLE Regulatory protein NOX01-delta
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1144)
AUTHORS Cheng, G. and Lambeth, J.D.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2002) Pathology, Emory University School of Medicine, 615 Michael Street, Atlanta, GA 30322, USA
FEATURES
1..1144
source
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.3"
1..1144
/gene="NOX01"
5..1132
/gene="NOX01"
/note="similar to p4phox; alternatively spliced"
/codon_start=1
/product="regulatory protein NOX01-delta"
/protein_id="AA038665.1"
/db_xref="GI:28394746"
/translation="MAGPRYPVSVOGAALVQIKELQTFAPSVRMSDGSDFVRSWDE
FRQKTKETFPVPEAGLLRRSDRVLPKLLGASLDAPLGRVGTSGRLARLQLEITY
SRLLIATAERVARSPFTITFAFQDLDEPALPESRVLPTPEEQPLSRAAGRLSIH
SLEAQSLRCIQPFCTQDRDRPFOAQESDILLRHPSGMLVENDRODTAMFPAPY
LEEAQCGREGGSLGSGSPQFCASRAYSSRLADELSVPAGARVRLVLETDQGMWLC
RYDRAGLLPAVLLRPEGLGALLSGTGRGDDDPAGEARGPPEPSQATAPPTVPTRP
SPGAIQSRSCCTVTTRALLRRPRQGRPRGCVDSVPHPPTTEQ"

ORIGIN
Query Match 96.4%; Score 1091; DB 9; Length 1144;
Best Local Similarity 98.4%; Pred. No. 4.1e-170;
Matches 1129; Conservative 0; Mismatches 0; Indels 18; Gaps 2;

QY 1 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGACCCCTGTGCAGATCAA 60
| | | | |
Db 1 AGCCATGGCAGGCCCCGATACCCAGTTTCAGTGCAGGGGCGACCCCTGTGCAGATCAA 60
| | | | |
QY 61 GAGGCTCCAAACGTTTGCCTTCTGTGCGCTGTGCAGCGCAGCGACACCTTCGTGCG 120
| | | | |
Db 61 GAGGCTCCAAACGTTTGCCTTCTGTGCGCTGTGCAGCGCAGCGACACCTTCGTGCG 120
| | | | |
QY 121 CAGGAGTTGGACGAAATTCAGGCGAGCTCAAGAAGACCTCAAGAGAGCTTCCCGGTGGA 180
| | | | |
Db 121 CAGGAGTTGGACGAAATTCAGGCGAGCTC--AAGACCTCAAGGAGACCTTCCCGGTGGA 177
| | | | |
QY 181 GCGCGGCTGTGCGGAGATGACCGGCTTCTCCAAAGCTTCTC----- 226
| | | | |
Db 178 GCGCGGCTGTGCGGAGATGACCGGCTTCTCCAAAGCTTCTC----- 237
| | | | |
QY 227 -GATGCACCACTGTTGGGACGCGTGGGCGCGACGAGCGCGGCTTGGCGCGCTGCAGCT 285
| | | | |
Db 238 GGATGCACCACTGTTGGGACGCGTGGGCGCGACGAGCGCGCTTGGCGCGCTGCAGCT 297
| | | | |
QY 286 GTTGGAAACCTATTTCTCGAGGCTGTGGCGACTGACAGAGCGCTGGCAACCGAC 345
| | | | |
Db 298 GTTGGAAACCTATTTCTCGAGGCTGTGGCGACTGACAGAGCGCTGGCAACCGAC 357
| | | | |
QY 346 GATCACTGGCTTCTTGGCAGCGCAACCCCTGGAGCTGGAGCCCGGCTGCCACCGGCGAG 405
| | | | |
Db 358 GATCACTGGCTTCTTGGCAGCGCAACCCCTGGAGCTGGAGCCCGGCTGCCACCGGCGAG 417
| | | | |
QY 406 CCGGCTGATCTCGCCACCCCGAGAGGAGCAGCTTCTTCTCGCGCTGCGGCGCGCTCTC 465
| | | | |
Db 418 CCGGCTGATCTCGCCACCCCGAGAGGAGCAGCTTCTTCTCGCGCTGCGGCGCGCTCTC 477
| | | | |
QY 466 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGAGAC 525
| | | | |
Db 478 CATCCACAGTCTGAGGCTCAGAGCTGCGCTGCTGAGCCCTTCTGTATCCAGGAGAC 537
| | | | |
QY 526 GCGGATAGGCTTTTCAGGCGCAGGCGCGAGGAGGCTGAGCTGTGCTGCGGCGACCC 585
| | | | |
Db 538 GCGGATAGGCTTTTCAGGCGCAGGCGCGAGGAGGCTGAGCTGTGCTGCGGCGACCC 597
| | | | |
QY 586 CTCAGGCTGTGGCTGTGGAGAAACGAAGACCGCGAGACCCCTGTGTTTCCAGGCGCCTA 645
| | | | |
Db 598 CTCAGGCTGTGGCTGTGGAGAAACGAAGACCGCGAGACCCCTGTGTTTCCAGGCGCCTA 657
| | | | |
QY 646 CCTCGAGGAGCGCGCCCGGCGCCAAAGCGCGGAGAGGAGCCCGTCCCTTAGGAGGAGCGG 705
| | | | |
Db 658 CCTCGAGGAGCGCGCCCGGCGCCAAAGCGCGGAGAGGAGCCCGTCCCTTAGGAGGAGCGG 717
| | | | |

QY	706	TCCCCAGTTCTGCTTCCCGGCTTACGAGAGCAGCCGCGCAGATGAGTGTCCGCTGCC	765
Db	718	TCCCCAGTTCTGCTTCCCGGCTTACGAGAGCAGCCGCGCAGATGAGTGTCCGCTGCC	777
QY	766	CGCGGGGGCGGCTGGCGGTGTGGAAACGTACAGACCGCGGCTGTGGCTATGACAGTA	825
Db	778	CGCGGGGGCGGCTGGCGGTGTGGAAACGTACAGACCGCGGCTGTGGCTATGACAGTA	837
QY	826	CGCGGACCGGGCGGCTTACTCCCGCGGTCTGCTGCGGCCCGAAGAGGCTGGCGCTCT	885
Db	838	CGCGGACCGGGCGGCTTACTCCCGCGGTCTGCTGCGGCCCGAAGAGGCTGGCGCTCT	897
QY	886	CTTGAGCGGACCGGGTTCCTGCGAGAGACGACCCCGCGGGTGAAGCCCGGGCTTCC	945
Db	898	CTTGAGCGGACCGGGTTCCTGCGAGAGACGACCCCGCGGGTGAAGCCCGGGCTTCC	957
QY	946	TGAACCTTCCAGCCACCGCCCTCCCGCCACCGTGTCCACCGACCTTCGCGGGCGC	1005
Db	958	TGAACCTTCCAGCCACCGCCCTCCCGCCACCGTGTCCACCGACCTTCGCGGGCGC	1017
QY	1006	CATCCAGAGCGCTGTGTCACCGTCAACGACGCGGCTTGAGCGCGCCACGCGGCCA	1065
Db	1018	CATCCAGAGCGCTGTGTCACCGTCAACGACGCGGCTTGAGCGCGCCACGCGGCCA	1077
QY	1066	GGCGCCCTCGAGGGTGTGAGTCTGTGCGCCACCCCGACGAGCAGTGAGCGCG	1125
Db	1078	GGCGCCCTCGAGGGTGTGAGTCTGTGCGCCACCCCGACGAGCAGTGAGCGCG	1137
QY	1126	AGGATCC 1132	
Db	1138	AGGATCC 1144	
RESULT 9			
LOCUS	BC019525		
DEFINITION	Mus musculus NADPH oxidase organizer 1, mRNA (cdna clone MGC:28618 IMAGE:4220018), complete cds.		
VERSION	BC019525		
KEYWORDS	BC019525.1 GI:18043495		
SOURCE	MGC.		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1260)		
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schett,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raja,S., Bosak,S.A., McEwan,P.J., Abramson,R.D., Mullaly,S.J., Bosak,S., Loquellano,N.A., Peters,G.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Kravinsky,M.I., Skalska,U., Smallos,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 1260)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		

JOURNAL

REMARK COMMENT

Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 37 Row: j Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 26080419.

FEATURES

source

1..1260

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:28618 IMAGE:4220018"

/tissue_type="Colon, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Co24"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1..1260

/gene="Noxol"

/note="synonyms: Snx28, P41NOX, P41NOXA, P41NOXB, P41NOXC"

/db_xref="LocusID:71893"

/db_xref="MGI:1919143"

27..1076

/gene="Noxol"

/codon_start=1

/product="Noxol protein"

/protein_id="AAH19525.1"

/db_xref="GI:18043496"

/db_xref="LocusID:71893"

/db_xref="MGI:1919143"

/translation="MSPRPVPSAHVALVQMDRLQTFAFVSCWSDNSDTFVRRSWDE FRLQKTKTTPFVEAGLLRSEQVLKPLDAPLLTRRGTGRGLVRLRLDTYQAL LATSHILRSSALHGFEVFKPDLPLPGLSVILPTPEPLSQPRGLDTHSLBAQ STPCQPHLDIRDPETHKAQEILDLRHPSGWLAVENKQDQVAFPPAPVLEVA TCQGESGLALQSGRQCTTQAYGSRDELVSFSGARVHVLSTDRGWMLCRYNGR TGLLPAMSLQPEGLSLGLRPGFSDSAGDKVAEDRTIPFVVVTRPCMSAIOSRCSITRRALAEQGTVRP"

93..389

/gene="Noxol"

/note="PX; Region: PX domain. PX domains bind to phosphoinositides"

/db_xref="CDD:pfam00787"

576..677

/gene="Noxol"

/note="SH3; Region: Src homology 3 domains"

/db_xref="CDD:smart00326"

747..887

/gene="Noxol"

/note="SH3; Region: Src homology 3 domains"

/db_xref="CDD:smart00326"

gene

CDS

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 46.5%; Score 526; DB 10; Length 1260;

Best Local Similarity 69.6%; Pred. No. 4, 7e-77;

Matches 753; Conservative 0; Mismatches 305; Indels 24; Gaps 2;

QY	1	AGCCATGGCAGGCCCCCGATACCCAGTTTCAGTGCAAGGGCAGCCCTGGTGCAGATCAA	60
DB	23	AGTCATGGCAAGCCCAAGACACCCAGTATCAGCCATGCTGTAGCTTGGTGCAATGGA	82
QY	61	GAGGCTCCAAAGTTTGCTTCTCTGTGGCTGGTTCAGACGGCAGCGACACTTGGTGGC	120
DB	83	CCGACTCCAGACATTTGGCTTCTCGTGTGCTGGTTCAGACAAACAGTGCACATTTGTATG	142
QY	121	CAGGATTTGGGACGAATTCAGGCAGCTCAAGAAGACCCCTCAAGGAGACCTTCCCAGTGA	180
DB	143	GAGGAGCTGGGATGAGTTTCAGGCAGCTCCAGAAGACCCCTTAAGAAACCTTCCCAGTGA	202
QY	181	GGCGGCTGTCTGCGGAGATCTGAACGGGTTCTCCAAAGCTTTCGATGTCACCACTGTT	240
DB	203	GGCAGGCTGTCTACGGAGATCTGAACAAGATTCTTCCCAAGCTTCTCGATGCTCCATTTGT	262
QY	241	GGGAGCGTGGGGCGCACGAGCCGGCGCTGCGCGCTGCAGCTGTTCGAGCTTTGGAAACCTATTC	300
DB	263	GACACGTGGGGGCGATCTGTGTGAGGACTGTGTACGTTTGGGCTGTCTGTGAGACCTATGT	322
QY	301	TCGGAGGCTGTGGCGACTTCGACGAGCGGCTGSCACGGAGCCCGACGATCACTGGGCTTCTT	360
DB	323	ACGGCATTTGTGGCAACCTCAGAACACATATTGAGGAGTTTCAGCACTTCACGGCTTCTT	382
QY	361	CGCACCGCAACCCCTGGACCTTGAGCCCGCGTGCACCCGGCAGCCGGGTGATCCTGCGC	420
DB	383	TGTATCCCAAACTCTGGATCTGGAGCCCATGTGCTCTCTGGCAGCTTGGTGTATCCTGCC	442
QY	421	CACCCACAGAGACGAGCCTCTTTCTCGCGCTGCGGGCGCGCTCTCCATCCACAGTCTCGGA	480
DB	443	TACACCAAGG---AGCCCTTATCCCAACCCAGAGGACGCTTGTACATTCACAGCCTCGGA	499
QY	481	GGCTCAGAGCTGCGCTGCGCTTCGACCCCTTCTGTACCCAGGACACGCGGATAGGCTTTT	540
DB	500	GGCTCAGAGCATTCCTCTGTGTACAGCCCTTCCACACTCTTGACATAAGACAGACAGCTTT	559
QY	541	TCAGGCGGAGCCGACGAGAGCGCTGGAAGTGTCTGTGCGGACACCCCTCAGGCTGTGCT	600
DB	560	CCACACCAAGGCTCAAGAAATTTCTGGACATATTACTAGCAGATCCTTCAGGCTGTGCT	619
QY	601	GGTGGAGAACGAAGACCGGACAGACCGCTGTGTTTCCAGCGCTTACTTGGAGGAGCGCGC	660
DB	620	GGTGGAGAACGAAGATCAGCAGGTAGCTGTGTTTCCAGCTCCCTTACTTGGAGAGGTAGC	679
QY	661	CCCGGCGCAAGGCGGGAGGAGGCCGTCTCCTAGGGAGCAGCGGTCCCCAGTTTGTGTC	720
DB	680	AACTGTCGAAGCCAGGAGTCAGGCCTGGCTTTGCAAGGAAGTGGGAGGCGATTTCTGCAC	739
QY	721	TTCCCGCGCTTACAGAGCAGCCGGGACAGATGAGCTGTCTGTGCCCCGGGGGCGCGGT	780
DB	740	TACCAGGCTTACAGGAGCATGCTCTGTATGAGCTATCCGTGTGCCCTCAGGGGACGCTGT	799
QY	781	CGCGCTGTTGGAAACGTTCAGACCGCGGTGTGTGCTATGCAGGTACGGCGACCGGGCGGG	840
DB	800	CCATGTCTGGAGACCTTCAGACCAAGGCTGTGTGCTGTGCAGGTATTAATGCGCCGACAGG	859
QY	841	CTTACTCCCGCGGTGTGTGTGCGGCGGAAGGGCTGGCGCTCTCTGTAGCGGGAACGGG	900
DB	860	CCTACTCCCTGCAATGTCTGTGCAACCTGAAGGGCTGGGCTCACTCTCTGGGACGCGAGG	919
QY	901	GTTCGTTGGAGGAGACGACCGCGGGGTGAGCCCGGGGCTTCCCTGNAACCTTCCCAGGC	960
DB	920	GTTCACAGACAGTCTGGGGCAGACAAGGTGG-----CTGAGGA	958
QY	961	CACCGCCCTCCCCCAACCGGTGCCCAACCCGACCTTTGCGCGGGCGCCATTCAGAGCCGCTG	1020
DB	959	CAGAAACATTTCCCTCTGTAGTACCACTCGTCCCTGTATGATGCTGCATTCAGAGTCTGATG	1018
QY	1021	CTGCAACGCTCACACGACGGGCCCTTGGAGCGGCGCCACAGGCGCCAGGGCGCGCTTCGAGG	1080
DB	1019	CTGCTCATTATCCGCGAGGACCTGGGACAGGAAACAGGGAATCTCGGGTTCCCGTTGAGG	1078

QY		1081 GT 1082 	
D6		1079 AT 1080	
RESULT 10			
AF539797			
LOCUS		1050 bp mRNA linear	ROD 03-FEB-2003
DEFINITION	Mus musculus NADPH oxidase organizer 1 (Noxol) mRNA,		complete cds.
ACCESSION	AF539797		
VERSION	AF539797.1 GI:25573151		
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus (house mouse)		
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;			
1 (bases 1 to 1050)			
AUTHORS	Banfi,B., Clark,R.A., Steger,K. and Krause,K.-H.		
TITLE	Two Novel Proteins Activate Superoxide Generation by the NADPH Oxidase NOX1		
JOURNAL	J. Biol. Chem. 278 (6), 3510-3513 (2003)		
PUBMED	12473664		
REFERENCE	2 (bases 1 to 1050)		
AUTHORS	Banfi,B., Clark,R.A., Steger,K. and Krause,K.-H.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-AUG-2002) Dept. of Geriatrics, University Hospitals of Geneva, Chemin du Petit-Bel-Air 2, Geneva CH-1225, Switzerland		
FEATURES	Location/Qualifiers		
source	. .1..1050 /organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="BALB/c"		
	/db_xref="taxon:10090"		
gene	. .1..1050 /gene="Noxol1"		
CDS	. .1..1050 /gene="Noxol1"		
	/note="p39NOX"		
	/codon_start=1		
	/product="NADPH oxidase organizer 1"		
	/protein_id="AA075142.1"		
	/db_xref="GI:25573152"		
	/translation="MASPRHPVSAHAVALVMQDLRQTFAFSVCWSNSDTFVRRSWDE FROLKTKTTFVEALGRREOVLPLKPDPAPLLTRRGTHGIVRLRLDTTVYAL LATSEHLRSSALHGFEVPKLDLEPMLPGSLVILPTPEEPLSQPSGSGLDIHSLEAQ SVCOPFTLDIRDRPFHTKAQLDILLRHSPSHWLVENKDQQVAWFPAYLEBEA TICQSQNALGGSGHQFCITTAQEGSRDELSSVPSGARHVLETSDRGWIICRYNGR TGLLPMSIQEPGLSGLRGPFPFDPSAGADKVAEDRTIPPVVTPTRPCMSAIOSRCCSI TRRALQGEOGTVRPR"		
ORIGIN			
Query Match	45.9%; Score 519.6; DB 10; Length 1050;		
Best Local Similarity	69.5%; Pred.No. 5.6e-76;		
Matches	746; Conservative 0; Mismatches 304; Indels 24; Gaps 2;		
Qy	5	ATGCAGGCCCGCCGATACCAGTTTCAGTGCAAGGGCAGCCCTGGTCAGATCAAGAGG 64	
D6	1	ATGCAGAAGCCAAGAACCCAGTATCAGCCCATTGTAGCCCTTGTGCAAAATGGACCGA 60	
Qy	65	CTCCAAAACGTTTTGCCCTTCTCTGTGCSCCTGGTCAGCGCAGCGCACACTTCGTGGCGCAGG 124	
D6	61	CTCCAGACATTTGCCCTTCTCCGTGTGCTGTCAGACAACAGTGACACATTTGTGGCGAGG 120	
Qy	125	AGTTGGGACGAATTACGGCAGTCAAGAAGACCTCAAGAGACCTTCCCGGTGGAGCGG 184	
D6	121	AGCTGGGATGAGTTTCAGGCAGCTCCCAAGAGACCTTTAAGAAAAACCTTCCAGTGGAGGCA 180	
Qy	185	GGCTCTGCGGAGATCTCAGCGCGTTCTCCCAAAGCTTCTCGATGCACCATGTTGGGA 244	
D6	181	GGCTCTGCTACGAGATCTGAAACAAGTTCTTCCCAAGCTTCTCGATGCTCATTTGCTGACA 240	
Qy	245	CGCGTGGGGCGCACGAGCGCGCCTTGGCGCGCTGACGCTGTGTGAAACCTATTCTCGG 304	

Db	241	CGTCGGGGGCACTACTGGTCAGAGACTGGTACGTTTGGGGCTGCTGGACACCTATGTACAG	300
Qy	305	AGGCTGTGCGGACTGCAGAGCGCGTGGCAGCGAGCGCCGACGATCACTGGCTTCTTCGCA	364
Db	301	GCATTGCTGCGCACTCAGAACACATATTGAGGAGTTCAGCACTTCACGGCTTCTTTGTA	360
Qy	365	CCGCAACCCCTGACCTGAGAGCGCGGCTGCCACCGCGCAGCGGGGTATCTTGCCCAAC	424
Db	361	CCCAACCTCTGATCTGGAGCGCCATGCTCCTCTGGCAGCCTGGTATCTCGCTCTACA	420
Qy	425	CCAGAGAGAGAGCTCTTTCTCGGCTGCGGGCGCTCTCCATCCACAGCTCTGAGGCT	484
Db	421	CCAGAGG---AGCCCTTATCCCAACCCAGAGGAGCCTTGACATTCATAGCCTGAGGCT	477
Qy	485	CAGAGCTGCGCTCTGTCAGCCCTTCTGTACCCAGGACACGCGGATAGGCTTTTCAG	544
Db	478	CAGAGCATTCCTGTGTACAGCCTTTCCACACTCTTGACATNAGACAGACCTTTCCAC	537
Qy	545	GCGCAGGCCAGAGAGACCTGAGAGTGTCTGCTCGGGACCCCTCAGGCTGGTGGTG	604
Db	538	ACCAAGGCTCAAGAAATTTCTGGACATATTACTACGACATCTTCAGGCTGGTGGTG	597
Qy	605	GAGAACGAGACCGGAGACCGCTGTGTTTCCAGCGCCCTACCTGAGGAGGGCGCCCG	664
Db	598	GAGAACAGATCAGCAGGTAGCTGTGTTTCCAGCTCCCTTACTCTGAGGAGGTAGCAACG	657
Qy	665	GGCAAGCGCGGAGGAGCGCCGTCTCCTAGGAGCAGCGGTCCCGATTTCTGTCTTC	724
Db	658	TGCCAAGGCCAGAGTCAGGCTGGCTTTGCAAGGAAGTGGAGGACAGTTCGCACTACT	717
Qy	725	CGGCTCTAGAGAGACCGCGCAGATGAGCTGTCTGTCGCCCGGGGGCGCGTGGC	784
Db	718	CAGGCTTACGAGGCGAGTCTGCTGATGAGTATCCGTGCTCCCTCAGGGGACGCTGCCAT	777
Qy	785	GTGTTGGAACGTTCAGACCGCGCTGTGCTGATGAGTACGGCAGCGGCGGCGCTTA	844
Db	778	GTCTGGAGACCTTCAGACCGAGCTGTGTGCTGTGCAAGTATTAATGGCCGGAACAGCCCTA	837
Qy	845	CTCCCGCGGTGTCTGTGCGCGCGGAAGGGCTGGGGCTCTCTGAGCGGAGCGGGTTC	904
Db	838	CTCCCTGCAATGTCGTGCAACCTGAAGGGCTGGGCTCGCTCTGGGAGGCCAGGGTTC	897
Qy	905	CGTGGAGAGACGACCCCGGGGTGAGCGCCGGGGTTCCTTGAACCTCCAGGCCAAC	964
Db	898	CCAGACAGTGTGGGCGAGACAAAGTGG-----CTGAGGACAGG	936
Qy	965	GCCCTCCCCACACCTGCCCCACCCGACCTTCGCCGGGCGCATCCAGAGCCCTGCTGC	1024
Db	937	ACCATTCCTCCCTGTAGTACCACTCGTCCCTGTATGAGTGCCCATCCAGAGTCGATGCTGC	996
Qy	1025	ACCGTCACAGCAGGCGCTTGGAGCGCGGCCACCGCGCGCCAGGGCGCCCTCGA	1078
Db	997	TCCATTACCGCAGGGCACTGGGACAGGAACAGGAGTCTGGGTTCCCGTTGA	1050
RESULT 11			
AB105906			
LOCUS			
DEFINITION			
Cavia porcellus NOXO1 mRNA for NADPH oxidase organizer 1, partial cds.			
AB105906			
VERSION			
AB105906.1 GI:42794029			
KEYWORDS			
SOURCE			
ORGANISM			
Cavia porcellus (domestic guinea pig)			
Cavia porcellus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.			
REFERENCE			
1			
Kawahara, T., Kuwano, Y., Teshima-Kondo, S., Takeya, R., Sumimoto, H., Kishi, K., Tsunawaki, S., Hirayama, T. and Rokutan, K.			
Role of Nicotinamide Adenine Dinucleotide Phosphate Oxidase 1 in Oxidative Burst Response to Toll-Like Receptor 5 Signaling in Large Intestinal Epithelial Cells			

JOURNAL	J. Immunol.	172 (5),	3051-3058	(2004)
PUBMED	14978110			
REFERENCE	2 (bases 1 to 759)			
AUTHORS	Kawahara, T. and Rokutan, K.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-MAR-2003) Kazuhito Rokutan, University of Tokushima, Department of Nutritional Physiology, School of Medicine; 3-18-15 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail:rokutan@nutr.med.tokushima-u.ac.jp, Tel:81-88-633-9246, Fax:81-88-633-7086)			
FEATURES				
source				
1..759				
/organism="Cavia porcellus"				
/mol_type="mRNA"				
/strain="Hartley"				
/db_xref="taxon:10141"				
1..759				
/gene="NOXO1"				
<1..>759				
/gene="NOXO1"				
/note="p41NOX"				
/codon_start=1				
/product="NADPH oxidase organizer 1"				
/protein_id="BAD11766.1"				
/db_xref="GI:42794030"				
/translation="WSDGSDTFCVRSWDEFRLQKTLKENFPFVEAGLLRRSDRLPKL DPAPLLSRGRTGRGLVRLRLNLTVOELLANVQVSWSPVLTSPFEPNRDILESALP PSMVLTLPVPEKPSRPVNSLDIHSLEIOSLCCVHPHTQDGRPFHVKAESLDVL LRHPGWLAVENEGQQKAPFAPYALEVPEGLMLQNSCTVFCASRAYKGSRADELSV PAGARVRVLETSDRGMWLCRYCDQGLLPFVLLQP"				
ORIGIN				
Query Match 36.7%; Score 415.4; DB 10; Length 759;				
Best Local Similarity 73.1%; Pred. No. 9e-59;				
Matches 568; Conservative 0; Mismatches 191; Indels 18; Gaps 2;				
Qy	92	TGTCAGACGGCAGCAGACCTTTCGTGCGCAGAGAGTTGGAGCAATTCAGGACGCTCAAG	151	
Db	1	TGTCGGATGGCAGTGATACCTTTGTTGCAGAACTGGGATGAGTTTCAGGACGCTCCAG	60	
Qy	152	AAGACCTTCAAGGAGACCTTCCCGGTGGAGCGGGCTGCTCGGAGATCTGACCGCGTT	211	
Db	61	AAACCCCTCAAGGAGAACTTTCAGTGGAGGCGAGGTCTGTGAGGAGATCTGTATCGCCT	120	
Qy	212	CTCCCAAAAGCTTCTCATGCACACCTGTGTGGACGCGTGGGGCGCACGAGCCCGCGCTG	271	
Db	121	CTCCCAAGCTTCCAGATGCACTTGTCTTCCGTGGAGGGCGCACAGTCTGTGCTCTG	180	
Qy	272	GCGCGCTCGAGCTGTGAAACCTTATCTCGAGGCTGCTGGCGACTGACAGCGCGTG	331	
Db	181	GTGCGTCTGAGCTGTGAAACCTTATGTGCAGGAACCTGCTAGCCAATGTGAACCAAGTG	240	
Qy	332	GCAGGAGCCCAAGATCACTGGCTTCTTGCACCGCAACCCCTGGACCTGGAGCCCGCG	391	
Db	241	TCATGGAGCCCAAGTGTCTTCAAGCTTCTTTAGCCCTCGAAACCAAGACCTGGAATCTGA	300	
Qy	392	CTGCCACCCCGCAGCGCGGTGATCTTCCCAACCCCAAGAGGAGCAGCTCTTTCTCGCGCT	451	
Db	301	CTGCTCCCAAGCAGCATGTGTGATCTGCTCTCCCAAGAGCCCTCAC---CTCGCCT	357	
Qy	452	GCGGGCCGCTCTCCATCCACAGTCTGGAGGCTTCAGAGCTTGGCTGCTGACGCCCTTC	511	
Db	358	GTGAACAGCCTTGTATATCCACAGCTTAGAGATTCAGAGCTCTGTTGCGTGCATCTTTC	417	
Qy	512	TGTACCCAGGACACCGCGGATAGGCTTTTCAGGCGCAGCGCCCAAGAGCCTTGGACGTG	571	
Db	418	CACACCCAGGACACACAGGAAAGGCCCTTCCATGTGAAGGCCCAAGAAAGCCTTGGATGA	477	
Qy	572	CTGCTGCGCACCCCTTCAGGCTGTGGCTGTGGAGAACCAAGACCGGACAGCCGCTTG	631	
Db	478	CTGCTACGACACCCCTTCAGGCTGTGGCTGTGGAGATGAAGGCCAGCAGAGGCTTG	537	
Qy	632	TTTCCAGCGCCCTTACCTGGAGGAGCGCGCCCGCGGCAAGGCGGAGGAGGCGCCGCTCC	691	

```
Db 538 TTTCCAGCTCCCTATCTGGAGTGGC-----CCAGAGGTGGCTTGATG 582
Qy 692 CTAGGAGCAGCGGTCCCGAGTTCTGTGTTCTCCCGGCTACGAGAGCAGCGCGCAGAT 751
Db 583 CTGCAAAAACAGTGGGAGTGTGTCTGTCTCCGAGCCTACAGGCGCAGCGTGTCTGAT 642
Qy 752 GAGCTGTCCGTCGCCCGCGCGCGCGTGGTGGTGGAACTCTCAGACCGCGGCTGG 811
Db 643 GAACCTGTAGTCCAGCGGAGCAGAGTGGGGTCTTGGAACCTTCAGACCGTGGATGG 702
Qy 812 TGGCTATGAGGTACGCGACCGCGGCGGCTACTCCCGCGGTGTGTTGGCGCG 868
Db 703 TGGCTGTGACAGTACTGTGACCAAGATGCTGTGCTTCACTGCTGCTGCAACG 759

RESULT 12
CQ730806
LOCUS CQ730806 412 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16740 from Patent WO02068579.
ACCESSION CQ730806
VERSION CQ730806.1 GI:42305651
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16740 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
location/Qualifiers
1..412
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 36.3%; Score 410.4; DB 6; Length 412;
Best Local Similarity 99.8%; Pred. No. 7,1e-58;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 705 GTCCCAAGTTCTGCTTCCCGGCTACGAGAGCAGCGCGCAGATGAGTGTCCGTGC 764
Db 1 GTCCCAAGTTCTGCTTCCCGGCTACGAGAGCAGCGCGCAGATGAGTGTCCGTGC 60
Qy 765 CCGCGGGGCGCGCGTGGTGGAAAGTCAAGCGCGGCTGGTGGCTATGCGAGT 824
Db 61 CCGCGGGGCGCGCGTGGTGGAAAGTCAAGCGCGGCTGGTGGCTATGCGAGT 120
Qy 825 ACGGCGACCGGGCGGGCTACTCCCGGCTGCTGCGCGCGGCGGAGGGCTGGCGCTC 884
Db 121 ACGGCGACCGGGCGGGCTACTCCCGGCGGCTGCTGCGCGCGGAGGGCTGGCGCTC 180
Qy 885 TCCTGAGCGGAGCGGGTTCGTTGAGGAGACGACCCCGCGGGTGAAGCCCGGGGCTTCC 944
Db 181 TCCTGAGCGGAGCGGGTTCGTTGAGGAGACGACCCCGCGGGTGAAGCCCGGGGCTTCC 240
Qy 945 CTGAACCTTCCCGAGGCAACCGCTCCCGGCGGCTGGTGGCGGCTGGCGGGG 1004
Db 241 CTGAACCTTCCCGAGGCAACCGCTCCCGGCGGCTGGTGGCGGCTGGCGGGG 300
Qy 1005 CCATCCAGAGCGGCTGTGTCACCGTCAACGAGGGGCTTGAGCGGGCGCCACGGCGCC 1064
Db 301 CCATCCAGAGCGGCTGTGTCACCGTCAACGAGGGGCTTGAGCGGGCGCCACGGCGCC 360
Qy 1065 AGGGCGGCTTGAGGGTGGTGGAGCTCTGTGCGCGCACCCCGACGACGAGCA 1116
Db 361 AGGGCGGCTTGAGGGTGGTGGAGCTCTGTGCGCGCACCCCGACGACGAGCA 412
```

```
RESULT 13
AF399754
LOCUS AF399754 720 bp mRNA linear ROD 22-AUG-2001
DEFINITION Mus musculus SNX28 (Snx28) mRNA, complete cds.
ACCESSION AF399754
VERSION AF399754.1 GI:15277910
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Hong,W
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2001) Membrane Biology Laboratory, Institute of
Mol. & Cell. Biol., 30 Medical Drive, Singapore 117609
FEATURES
source
location/Qualifiers
1..720
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..720
/gene="Snx28"
1..720
/gene="Snx28"
/notes="member of sorting nexin family"
/codon_start=1
/product="SNX28"
/protein_id="AAK94017.1"
/db_xref="GI:15277911"
/translation="MASPRHPVSAHAVALVQMDRLQTFAPSVCHSDNSDTFVRSWDE
FRQKTLTKKTFPVEAGLLRRSEQLPAPLPLTRGHTGRGLVRLLRLDTIVQAL
LATSEHLRSALHGFVPKPLDLEPMLPPSLVILPTPEPLSPQPSGLDIHLSLEAQ
SIPCVQPFHTLDIRDPFHTKAQBILDILLRHPSSMWLVENKQQQVAFWPAPYLEEVA
TCQGESGLALQSGMSPSL"

ORIGIN
Query Match 33.5%; Score 378.8; DB 10; Length 720;
Best Local Similarity 72.2%; Pred. No. 9.7e-53;
Matches 507; Conservative 0; Mismatches 192; Indels 3; Gaps 1;
Qy 5 ATGCAGCGCCCCCATACCCAGTTTCAGTGCAGCGGCGAGCCCTGTCAGATCAAGAGG 64
Db 1 ATGCAGCGCCCCAAGACACCCAGTATCAGCCCATCTGTAGCTTGTGCAAAATGACCGA 60
Qy 65 CTCGAAAGCTTTGCTTCTCTGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 124
Db 61 CTCGAGACATTTGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy 125 AGTTGGGAGCAATTCAGGAGCTCAAGAGACCCCTCAAGGAGACCTTCCCGGTGAGGCG 184
Db 121 AGCTGGGATGAGTTTCAGGAGCTCCAGAGACCCCTTAAGAAAAACCTTCCAGTGGAGCA 180
Qy 185 GGCCTGTGCGGAGATCTGACCGGCTTCTCCCAAGCTTCTCGATGCAACCACTATTCTCGG 244
Db 181 GGCCTGTGACGAGATCTGAACAAAGTTCTTCCCAAGCTTCTCGATGCTCCATTGCTGAC 240
Qy 245 CGCGTGGGCGCACGAGCGCGGCTGTGCGGCTGTGAGCTGTGTGGAAACCTATTCTCGG 304
Db 241 CGTGGGGGCATCTGCTGAGGACTGGTACGTTTGGGCTGTGTGGACACCATATGTACAG 300
Qy 305 AGGCTGTGGGCACTGAGAGCGGTGGGCAAGAGCCCGAGCATCATCTGGCTTCTTCGCA 364
Db 301 GCATTTGTGGCAACTCAGAACACATATTGAGGAGTTTCAGCACTTTCAGGGCTTCTTTGTA 360
Qy 365 CCGCAACCCCTGGACCTGGAGCGGCTGCGCACCGCGAGCGGGTGTATCTGCGCCACC 424
Db 361 CCCAACTCTGGATCTGGAGCCCATGCTGCTCTGTGAGCGCTGGTGTATCTGCTTACA 420
Qy 425 CCAGAGGAGCAGCCTCTTTCTGCGGCTGTGCGGCGGCTCTCCATCCACAGTCTGAGGCT 484
```

Db	421	CCAGAGG---AGCCCTTATCCCAACCCAGAGGAGCGCTTGACATTCATAGCCTGGAGGCT	477	/map="16p13.3"		
Qy	485	CAGAGCCTGCGCTCGCTGACGCCCTTCTGTACCAGACACGCGGGATAGCCTTTTCAG	544	/clone="109-8C"		
Db	478	CAGAGCATTCCTGTGTACAGCCTTTCCACACTCTTGACATATAGACAGACCTTTCCAC	537	/note="GRAIL 2 excellent exon, frame 1"		
Qy	545	GCCAGGCCAGAGAGCCTGAGAGTCTGCTGCTGGGACCCCTCAGCTGGTGGTG	604	/rpt_family="LTR8"		
Db	538	ACCAAGGCTCAAGAAATTCGAGCATATTACTACGACATCTTCAAGCTGGTGGTG	597	/rpt_family="Alu"		
Qy	605	GAGAACGAGACCGGAGACCGCTGTGTTTCCAGCGCCCTACCTGGAGGAGCGGCCCG	664	/rpt_family="LTR8"		
Db	598	GAGAACAGGATCAGCAGGTAGCTGTGTTTCCAGCTCCCTACCTGGAGGAGTAGCAACG	657	/rpt_family="Alu"		
Qy	665	GGCCAAAGCGCGGAGGAGCGCGCTCCCTAGGAGCAGCGGT	706	/note="(A)20"		
Db	658	TGCCAAGGCCAGGAGTCAGGCGCTGGCTTTCAGAGGAAGTGT	699	/rpt_type=tandem		
RESULT 14				/rpt_unit="a"		
AC005606/c				4171..4477		
LOCUS				/rpt_family="Alu"		
DEFINITION	AC005606	80662 bp	DNA	linear	complement(4596..4715)	exon, frame 1"
ACCESSION	AC005606.3	GI:28144383			complement(4932..5027)	exon, frame 1"
VERSION					complement(5722..5823)	exon, frame 0"
KEYWORDS	HTG.				complement(6035..6195)	exon, frame 2"
SOURCE	Homo sapiens (human)				complement(6291..6477)	exon, frame 0"
ORGANISM	Homo sapiens				complement(7036..7609)	
REFERENCE					/rpt_family="Alu"	
AUTHORS	Ricke,D.O.				/note="(T)23"	
TITLE	Large Scale Sequence Analysis and Annotation with the Sequence				/rpt_type=tandem	
JOURNAL	Unpublished				/rpt_unit="t"	
REFERENCE	2 (bases 1 to 80662)				complement(7630..8255)	
AUTHORS	Ricke,D.O., Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E.,				/rpt_family="Alu"	
	Robinson,D., Jones,M., Buckingham,J., Chasteen,L., Thompson,S.,				/note="(T)22"	
	Goodwin,L., Bryant,J., Tesmer,J., Meinke,L., Longmire,J.,				/rpt_type=tandem	
	White,S., Ueng,S., Tatum,O., Campbell,C., Fawcett,J., Maltbie,M.,				/rpt_unit="t"	
	Mistra,M. and Deaven,L.				complement(8361..8633)	
TITLE	Sequencing of Human Chromosome 16p13.3				/rpt_family="Alu"	
JOURNAL	Unpublished				complement(9941..10076)	
REFERENCE	3 (bases 1 to 80662)				/note="GRAIL 2 excellent	exon, frame 2"
AUTHORS	Ricke,D.O., Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E.,				complement(10484..10748)	
	Robinson,D., Jones,M., Buckingham,J., Chasteen,L., Thompson,S.,				/rpt_family="Alu"	
	Goodwin,L., Bryant,J., Tesmer,J., Meinke,L., Longmire,J.,				10821..10868	
	White,S., Ueng,S., Tatum,O., Campbell,C., Fawcett,J., Maltbie,M.				/note="(ATT)15"	
	and Deaven,L.				/rpt_type=tandem	
TITLE	Direct Submission				/rpt_unit="att"	
JOURNAL	Submitted (04-SEP-1998) Center for Human Genome Studies, DOE Joint				complement(11060..11358)	
	Genome Institute, Los Alamos National Laboratory, MS M888, Los				/rpt_family="Alu"	
	Alamos, NM 87545, USA				complement(11490..11760)	
REFERENCE	4 (bases 1 to 80662)				/rpt_family="Alu"	
AUTHORS	Mundt,M.O.				complement(11971..12139)	
TITLE	Direct Submission				/note="GRAIL 2 excellent	exon, frame 1"
JOURNAL	Submitted (02-DEC-2000) Biosciences, Joint Genome Institute/ Los				complement(12322..12616)	
	Alamos National Laboratory, MS M888, Los Alamos, New Mexico 87545,				/rpt_family="Alu"	
	U.S.				complement(13053..13245)	
REFERENCE	5 (bases 1 to 80662)				/note="GRAIL 2 excellent	exon, frame 0"
AUTHORS	Mundt,M.O.				14232..14596	
TITLE	Direct Submission				/rpt_family="Alu"	
JOURNAL	Submitted (30-JAN-2003) Biosciences, Joint Genome Institute/ Los				/rpt_family="Alu"	
	Alamos National Laboratory, MS M888, Los Alamos, New Mexico 87545,				complement(15014..15182)	
	U.S.				complement(15261..15550)	
COMMENT	On Jan 30, 2003 this sequence version replaced gi:11527450.				/rpt_family="Alu"	
FEATURES	Location/Qualifiers				complement(15605..15894)	
source	1..80662				/rpt_family="Alu"	
	/organism="Homo sapiens"				16072..16359	
	/mol_type="genomic DNA"				/rpt_family="Alu"	
	/db_xref="taxon:9606"				complement(16472..17206)	
	/chromosome="16"					

```
misc_feature      /rpt family="LTR12"
18578..18845      /note="GRAIL 2 excellent exon, frame 1"
repeat_region    complement(19317..19602)
misc_feature      /rpt family="Alu"
20244..20383      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      complement(21204..21361)
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      complement(21588..21747)
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      complement(21812..22000)
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      complement(22240..22347)
repeat_region    /note="GRAIL 2 excellent exon, frame 0"
22427..22718      /rpt family="Alu"
misc_feature      complement(23337..23456)
repeat_region    /note="GRAIL 2 excellent exon, frame 1"
misc_feature      complement(25378..25789)
repeat_region    /rpt family="Alu"
26365..26510      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      26816..26959
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      27525..27980
repeat_region    /note="GRAIL 2 excellent exon, frame 2"
28148..29029      /rpt family="Alu"
repeat_region    complement(29102..29373)
repeat_region    /rpt family="Alu"
complement(29472..29999)
repeat_region    /rpt family="Alu"
complement(30309..30626)
repeat_region    /rpt family="Alu"
32251..32538      /rpt family="Alu"
misc_feature      33147..33198
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      33288..33383
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      33467..33514
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      33629..33770
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      33854..33942
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      34019..34189
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      34436..34529
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      34620..34801
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      34899..35040
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      35147..35205
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      35905..36052
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      36145..36375
misc_feature      /note="GRAIL 2 excellent exon, frame 0"
misc_feature      36447..36517
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      36605..36764
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      36833..36891
misc_feature      /note="GRAIL 2 excellent exon, frame 1"
misc_feature      37049..37152
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      37240..37328
misc_feature      /note="GRAIL 2 excellent exon, frame 2"
misc_feature      37404..37562
misc_feature      /note="GRAIL 2 excellent exon, frame 1"

misc_feature      complement(38516..38633)
/note="GRAIL 2 excellent exon, frame 1"

Query Match      30.4%; Score 344; DB 9; Length 80662;
Best Local Similarity 85.1%; Pred. No. 1.4e-47;
Matches 429; Conservative 0; Mismatches 0; Indels 75; Gaps 1;

QY 704 GGTCCCGAGTTCTGCTTCCGCGCCTTACGAGAGCAGCGCGGACGAGATGAGCTGTCGCTG 763
Db |||||||
QY 38634 GGTCCCGAGTTCTGCTTCCGCGCCTTACGAGAGCAGCGCGGACGAGATGAGCTGTCGCTG 38575
Db |||||||
QY 764 CCCGCGGGGCGCGCTGCGCGTGTGGAAACGTCAGACCGCGGCTGCTGGCTAT----- 818
Db |||||||
QY 38574 CCCGCGGGGCGCGCTGCGCGTGTGGAAACGTCAGACCGCGGCTGCTGGCTATGCAGG 38515
QY 819 -----
Db 38514 TACGCGGAGCGCGCTGGGCGAGGCGTACGCGAGCGGGCGGCGCCCTAACACAC 38455
QY 819 -----GCAGGTACGGCGAGCGGGCGGCGCTACTCCCGGGTCTGCTGGCGCG 868
Db 38454 CCCGCCCTTCGAGGTACGGCGACCGGGCGGCGCTACTCCCGGGTCTGCTGGCGCG 38395
QY 869 GAAGGGCTGGCGCTCTCTCTGAGCGGACGCGGGTTCCGTGAGGAGACGACCCGCGGGT 928
Db 38394 GAAGGGCTGGCGCTCTCTCTGAGCGGACGCGGGTTCCGTGAGGAGACGACCCGCGGGT 38335
QY 929 GAGCGCGGGGCTTCCCTGAACCTCCAGGCCACCGCGCTCCCGCCACCGTGCACCC 988
Db 38334 GAGCGCGGGGCTTCCCTGAACCTCCAGGCCACCGCGCTCCCGCCACCGTGCACCC 38275
QY 989 CGACCTTCCGCGGCGCATCCAGAGCGGCTGTCACCGTCAACGAGGCGCTGAG 1048
Db 38274 CGACCTTCCGCGGCGCATCCAGAGCGGCTGTCACCGTCAACGAGGCGCTGAG 38215
QY 1049 CGGCGCCACGCGCGCGCGCTCGAGGGCGCGCTCGAGGGTGGTGACTCTGTCCGCGACCC 1108
Db 38214 CGGCGCCACGCGCGCGCGCTCGAGGGCGCGCTCGAGGGTGGTGACTCTGTCCGCGACCC 38155
QY 1109 ACGGAGCAGTGAGCGCGAGGATCC 1132
Db 38154 ACGGAGCAGTGAGCGCGAGGATCC 38131

RESULT 15
AX834341/c      2922 bp      DNA      linear      PAT 15-DEC-2003
LOCUS      Sequence 1465 from Patent EP1347046.
DEFINITION      AX834341
ACCESSION      AX834341
VERSION      AX834341.1 GI:39920476
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.
TITLE      Full-length cDNA sequences
JOURNAL      Patent: EP 1347046-A 1465 24-SEP-2003;
FEATURES
source      Research Association for Biotechnology (JP)
1..2922
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match      30.2%; Score 342.4; DB 6; Length 2922;
Best Local Similarity 84.9%; Pred. No. 6.5e-47;
Matches 428; Conservative 0; Mismatches 1; Indels 75; Gaps 1;
```

```

Qy 704 GGTCCCAGTTCTGTGTTCTCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTG 763
Db 679 GGTCCCAGTTCTGTGTTCTCCGCGCTACGAGAGCAGCGCGCAGATGAGCTGTCCGTG 620
Qy 764 CCGCGGGGGCGCGCGGTGCGGTGTTGGAAACGTACAGACCGCGGCTGTGGCTAT----- 818
Db 619 CCGCGGGGGCGCGCGGTGCGGTGTTGGAAACGTACAGACCGCGGCTGTGGCTATGCAGG 560
Qy 819 ----- 818
Db 559 TACGCGGAGCGGGCTGGGCGAGGCGAGGCTAGCCGAGGCGGGGCGCTAACACAC 500
Qy 819 -----GCAGGTACGCGGACCGGGCGGGCTACTCCCGCGGTGTGTGCGGCCG 868
Db 499 CCGCGCCCTCGCAGGTACGCGGACCGGGCGGGCTACTCCCGCGGTGTGTGCGGCCG 440
Qy 869 GAAGGGCTGGCGCTCTCTGAGCGGGAAGGGGTTCCGTGGAGAGACGACCCCGGGGT 928
Db 439 GAAGGGCTGGCGCTCTCTGAGCGGGAAGGGGTTCCGTGGAGAGACGACCCCGGGGT 380
Qy 929 GAGGCCCGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTCCACCC 988
Db 379 GAGGCCCGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGTCCACCC 320
Qy 989 CGACCTTCGCGGGCGCCATCCAGAGCCGCTGTGCACCGTCAACGAGGGGCCCTGGAG 1048
Db 319 CGACCTTCGCGGGCGCCATCCAGAGCCGCTGTGCACCGTCAACGAGGGGCCCTGGAG 260
Qy 1049 CGGCGCCACGCGCGCGAGGGCGGCCCTCGAGGGTGTGTGACTCTGTGCGGCACCCACG 1108
Db 259 CGGCGCCACGCGCGCGAGGGCGGCCCTCGAGGGTGTGTGACTCTGTGCGGCACCCACG 200
Qy 1109 ACGGAGCAGTGAGCGGAGGATCC 1132
Db 199 ACGGAGCAGTGAGCGGAGGATCC 176

```

Search completed: May 30, 2005, 21:37:31
Job time : 8117 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 30, 2005, 18:26:34 ; Search time 5919 Seconds
(without alignments)
7279.735 Million cell updates/sec

Title: US-10-621-113-3
Perfect score: 1132
Sequence: 1 agccatggcagccccccgat.....agcagtgcgagcaggatcc 1132

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524.4	46.3	2195	AK088226	AK088226 Mus muscu
2	466.8	41.2	790	BG470266	BG470266 602533785
3	432.2	38.2	519	BX283317	BX283317 BX283317
C 4	429	37.9	599	CA311857	CA311857 UI-CP-FNO
C 5	427.4	37.8	751	AL042759	AL042759 DRFP434C
6	421.2	37.2	623	BE793330	BE793330 601588311
7	421	37.2	473	BG328824	BG328824 602427426
8	381.2	33.7	1571	AK009605	AK009605 Mus muscu
9	376.2	33.2	904	BG871968	BG871968 602793072
10	347	30.7	726	BI409553	BI409553 602963963
11	334.8	29.6	544	BE751524	BE751524 203692 MA
12	326.4	28.8	773	BG247058	BG247058 602359714
13	318.4	28.1	918	BF581002	BF581002 602100116
14	317.4	28.0	602	CB113007	CB113007 K-EST0154
15	309.2	27.3	967	BQ935073	BQ935073 AGENCOURT
16	304.4	26.9	584	AW743151	AW743151 up63D08.Y
17	300	26.5	635	AA543925	AA543925 vk40C12.I
18	282.4	24.9	984	BY762147	BY762147 BY762147
19	281.8	24.9	541	AW743167	AW743167 up63D12.Y
20	274.8	24.3	542	AA269948	AA269948 v364G11.I
21	268	23.7	274	AW163511	AW163511 au95F11.Y
22	258	22.8	865	CO810832	CO810832 AGENCOURT
23	242.6	21.4	743	BI083169	BI083169 602875260
24	233.8	20.7	559	AV594936	AV594936 AV594936

25	232	20.5	582	5	BP346589	BP346589
26	230	20.3	480	4	BI534701	BI534701 397823 MA
C 27	222	19.6	349	2	BF591178	BF591178 7h43a06.X
28	217.8	19.2	450	5	BQ265615	BQ265615 NTSC ff07
29	206.4	18.2	437	2	BE554337	BE554337 uF41C11.Y
C 30	189.8	16.8	645	7	CN284040	CN284040 170004705
C 31	185.8	16.4	245	1	AA576374	AA576374 nm75a09.B
32	185.8	16.4	288	2	AW358014	AW358014 41850 MAR
C 33	181.4	16.0	898	5	BQ222005	BQ222005 AGENCOURT
34	180.4	15.9	837	6	BY709278	BY709278 BY709278
C 35	180	15.9	303	1	AI435004	AI435004 th76G05.X
36	178.4	15.8	658	1	AL042758	AL042758 DRFP434C
37	173	15.3	405	1	AI019507	AI019507 ua89d10.I
38	165.8	14.6	333	1	AI173340	AI173340 uc29f05.I
39	164	14.5	572	2	BF140326	BF140326 601787835
40	157.8	13.9	612	4	BI113592	BI113592 602897862
41	154	13.6	733	4	BI415175	BI415175 602990692
42	152.2	13.4	906	5	BQ959588	BQ959588 AGENCOURT
C 43	149.6	13.2	306	2	AW083983	AW083983 xc36G09.X
44	148.6	13.1	427	2	BF522080	BF522080 UI-R-Y0-a
45	147.8	13.1	812	4	BI112184	BI112184 602899851

ALIGNMENTS

RESULT 1	AK088226	2195 bp	linear	HTC 03-APR-2004
LOCUS	Mus musculus	2 days neonate thymus	thymic cells	CDNA, RIKEN
DEFINITION	full-length enriched library, Clone:E430007K11	product:2310034C04RIK	PROTEIN (SNX28)	homolog [Mus musculus], full insert sequence.
ACCESSION	AK088226	1	GI:26353183	
VERSION	AK088226.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	Matches	752; Conservative	0; Mismatches	306; Indels	24; Gaps	2;
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	QY	1	AGCCATGTCAGGCCCCCGATACCCAGATTTTCAGTGCAGGGGAGCCCTGTGTGACAGATCAA	60		
JOURNAL	Nature 420, 563-573 (2002)	Db	970	AGTCATGGCAAGCCCAAGACACCCAGTATCAGCCCATCTGTAGCTTGGTGCAATGGA	1029		
REFERENCE	6 (bases 1 to 2195)	QY	61	GAGGCTCCAAAGCTTGGCTTCTCTGTGGCTGGTGCAGCGGACGACACCTTCGTGGG	120		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaiura,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.	Db	1030	CCGACTCCAGACATTTGGCTTCTCCGTGTGTGGTGCAGCAACAGTGCACATTTGTGG	1089		
TITLE	Direct Submission	QY	121	CAGGATTTGGGACGAATTCAGCGAGCTCAAGAAGACCCTCAAGGAGACTTCCCGGTGA	180		
JOURNAL	Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	Db	1090	GAGGAGTGGATGATTCAGCGAGCTCCAGAGACCTTAAGAAAACCTTCCCAAGTGA	1149		
		QY	181	GGCGGCTGTCTGGGAGATGTGACCGGTTCTCCAAAGCTTCTCGATGACACACTGTT	240		
		Db	1150	GGCAGGCTGTCTACGGAGATCTGAACAAGTTCTTCCCAAGCTTCTGATGTCTCAATTG	1209		
		QY	241	GGGACGCTGGGGCGCAGAGCCGCGCTCGCGCTGCAGCTGTGTGGAAACCTATTTC	300		
		Db	1210	GACAGTCGGGGGATACCTGTCAGGACTGGTGTTCGGCTGTGGACACCTATGT	1269		
		QY	301	TCGGAGGCTGTGGCGACTGTCAGAGCGCGTGGCAGCGCCGACGATCCTCTGCTT	360		
		Db	1270	ACAGGCATTTCTGGCAACCTCAGAACACATATTGAGGAGTTTCAACCTTCAGCGCT	1329		
		QY	361	CGCACCGCAACCTCTGACCTTGGAGCCCGCTGTCACCCCGGAGCCGGGTGATCTG	420		
		Db	1330	TGTATCCCAACCTCTGGATCTGGAGCCCATCTGCTCTCTGGCAGCTGTGATCCTGC	1389		
		QY	421	CACCCAGAGAGCAGCTCTTCTCGGCTCGCGGCTGCTCTCCATCCACAGTCTCGA	480		
		Db	1390	TACACCCAGAGG---AGCCCTATTCCCAACCCAGAGGAGCTTGACATTCATAGCT	1446		
		QY	481	GGCTCAGAGCTGCGCTGCTGACGCTTCTGTACCCAGACACGCGGATAGGCTTTT	540		
		Db	1447	GGCTCAGAGCATTCCTGTGTACAGCTTCCACACTTTGCACATAAGACAGACACT	1506		
		QY	541	TCAGGCGCAGGCCAGGAGACCTTGGAGTGTCTGTCGGGACCCCTCAGGCTGTGCT	600		
		Db	1507	CCACACCAAGGCTCAAGAAATTTCTGGACATATTACTAGACATCTTCAGGCTGGT	1566		
		QY	601	GGTCGAGAACGAACCGGACGCGCTGCTTCCAGCGCCCTACCTGGAGGAGCGCG	660		
		Db	1567	GGTCGAGAACGAAGATCAGCAAGTAGCTGTGTTCCAGCTCCCTACCTTGGAGAGT	1626		
		QY	661	CCCGGCGCAGCGCGGAGGAGGCGCTGCTTCCAGGAGCAGCGTCCCGAGTCTGTC	720		
		Db	1627	AACGTGCCAAGCCAGGAGTCAGGCTGCGCTTGCAGGAAGTGGGAGGAGTCTG	1686		
		QY	721	TTCCGCGCTTACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCCGGGGGCGCG	780		
		Db	1687	TACCCAGGCTTACGAGGAGTGTCTGTATGAGTATCCGTCGCTCAGGCGGACGT	1746		
		QY	781	GCGGCTGTGGAACCTCAGACCGCGCTGTGGCTATGACAGGTACGCGACCGGCG	840		
		Db	1747	CCATGTCTGGAGACCTCAGACCGAGGCTGTGGCTGTGAGGATTAATGCGCGG	1806		
		QY	841	CCTACTCCCGCGTGTCTGTCGCGCGGAAAGGCTGGGCGCTCTCTCTGAGCGG	900		
		Db	1807	CCTACTCCCTCAATGTCTGTCAACCTGAAGGGCTGGGCTCGCTCTCTGGG	1866		
		QY	901	GTTCTGTGGAGAGACGACCGCGGGGTGAGCGCGGGCTTCCCTGAAACCTCCAG	960		
		Db	1867	GTTCCAGACAGTGTCTGGGGCAGACAAGGTGG-----CTGAGGA	1905		
		QY	961	CACCGCCCTCCCGCCCGCTGCCACCCGACCTTTCGCGCGGCGCATCCAGAGCG	1020		
		Db	1906	CAGGACCATTTCCCTCTGTATGATACCACTGCTCTCTGTATGAGTGCATCCAG	1965		
		QY	1021	CTGCACCGTCAACGAGGCGCTTGGAGCGCGCGCCACGCGCCAGGCGCGCTCG	1080		
		Db	1966	CTGCTCCATTACCCGCGAGGCGCTGGGACAGGAACAAGGAGCTCGGGTTC	2025		
FEATURES	source	polyA_signal	2172..2177				
		polyA_site	2195				
ORIGIN	Query Match						
	Best Local Similarity	46.3%; Score 524.4; DB 3; Length 2195;					
		69.5%; Pred. No. 2.8e-98;					


```
Query Match      38.2%; Score 432.2; DB 5; Length 519;
Best Local Similarity 98.9%; Pred. No. 3.2e-79;
Matches 434; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 694 AGGAGACAGCGTCCCACTTCTGTGCTTCCCGCGCTACGAGAGAGCGCGCGCAGATGA 753
DB 17 AGGAGACAGCGTCCNTCCNTCTGTGCTTCCCGCGCTACGAGAGAGCGCGCGCAGATGA 76
QY 754 GCTGTCCGTGCGCGCGCGCGCGCGCTGTGCGCTGTGGAACGTCAGACCGCGCTGGTG 813
DB 77 GCTGTCCGTGCGCGCGCGCGCGCGCTGTGCGCTGTGGAACGTCAGACCGCGCTGGTG 136
QY 814 GCTATGACAGTACGCGACCGCGCGCGCTACTCCCGCGGTGCTGCGCGCGCGAGG 873
DB 137 GCTATGACAGTACGCGACCGCGCGCGCTACTCCCGCGGTGCTGCGCGCGCGAGG 196
QY 874 GCTGGCGCTCTCTGAGCGGACCGGGTTCGCTGGAGGAGACGACCCCGCGGGTGAGGC 933
DB 197 GCTGGCGCTCTCTGAGCGGACCGGGTTCGCTGGAGGAGACGACCCCGCGGGTGAGGC 256
QY 934 CGGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGCTGCCACCGGACC 993
DB 257 CGGGGCTTCCCTGAACCTCCAGGCCACCGCCCTCCCGCCACCGCTGCCACCGGACC 316
QY 994 TTCGCGCGGCGCATCCAGAGCGCTGCTGACCGCTCACACGACGCGCGCTGGAGCGCG 1053
DB 317 TTCGCGCGGCGCATCCAGAGCGCTGCTGACCGCTCACACGAGCGCGCTGGAGCGCG 376
QY 1054 CCACGCGCGCCAGGCGCGCTTCGAGGGTGCGTGGACTCTGTGCGCGACCCCGACGCGA 1113
DB 377 CCACGCGCGCCAGGCGCGCTTCGAGGGTGCGTGGACTCTGTGCGCGACCCCGACGCGA 436
QY 1114 GCAGTGAGCGCGAGGATCC 1132
DB 437 GCAGTGAGCGCGAGGATCC 455

RESULT 4
LOCUS      CA311857      599 bp      mRNA      linear      EST 04-NOV-2002
DEFINITION UI-CF-FN0-afe-d-10-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
ACCESSION  CA311857
VERSION     CA311857.1  GI:24529955
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 599)
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES      Ronaldo,M.F., Lennon,G. and Soares,M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.regen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.

FEATURES             Location/Qualifiers
     CDS             1..599
     ORIGIN
         Query Match      37.9%; Score 429; DB 6; Length 599;
         Best Local Similarity 100.0%; Pred. No. 1.5e-78;
         Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

         QY 704 GGTCCCGAGTTCTGTGCTTCCCGCGCTACGAGAGACGCGCGCAGATGAGCTGCCGTG 763
         DB 569 GGTCCCGAGTTCTGTGCTTCCCGCGCTACGAGAGACGCGCGCAGATGAGCTGCCGTG 510
         QY 764 CCCGCGGGCGCGCGTGCCTGTGGAACGTCAGACCGCGCTGTGGCTATGCAAGG 823
         DB 509 CCCGCGGGCGCGCGTGCCTGTGGAACGTCAGACCGCGCTGTGGCTATGCAAGG 450
         QY 824 TACGCGACCGGCGCGCTACTCCCGCGCTGTGCTGCGCGCGAAGGCTGGCGCT 883
         DB 449 TACGCGACCGGCGCGCTACTCCCGCGCTGTGCTGCGCGAAGGCTGGCGCT 390
         QY 884 CTCCTGAGCGGACGGGTTCCGTGGAGAGACGACCCCGCGGGTGAGCGCGGGCTTC 943
         DB 389 CTCCTGAGCGGACGGGTTCCGTGGAGAGACGACCCCGCGGGTGAGCGCGGGCTTC 330
         QY 944 CTGAACCTCCAGGCGCACCGCCCTCCCGCGCGTGCACCCGACCTTCGCGCGGCG 1003
         DB 329 CTGAACCTCCAGGCGCACCGCCCTCCCGCGCGTGCACCCGACCTTCGCGCGGCG 270
         QY 1004 GCCATCCAGACCGCTGCTGCACCGTCAACGAGGCGCTGGAGCGCGCGCCACGCGCGC 1063
         DB 269 GCCATCCAGACCGCTGCTGCACCGTCAACGAGGCGCTGGAGCGCGCGCCACGCGCGC 210
         QY 1064 CAGGCGCGCTCGAGGCTGCTGACTCTGTGCGGACCCCGACGAGGAGTGGAGG 1123
         DB 209 CAGGCGCGCTCGAGGCTGCTGACTCTGTGCGGACCCCGACGAGGAGTGGAGG 150
         QY 1124 CGAGGATCC 1132
         DB 149 CGAGGATCC 141

RESULT 5
LOCUS      AL042759/c      751 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFZp3434C0722.s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION  AL042759
VERSION     AL042759.2  GI:5935566
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 751)
AUTHORS     Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            POLYA=Yes.
```

TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 8, 1999 this sequence version replaced gi:5422206.
 Contact: MIPS

Ingolstaedter Landstr.1, D-95764 Neuherberg, Germany
 This is the 3' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University,
 Munich/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 xl sequence also available.
 This clone (DKFZp34C0722) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1..751
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp34C0722"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
 Query Match 37.8%; Score 427.4; DB 1; Length 751;
 Best Local Similarity 99.8%; Pred. No. 3.3e-78;
 Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 704 GGTCCCCAGTTCTGTTCCCGCGCTACGAGAGCAGCGCCAGATGATGTCGCTG 763
 Db 637 GGTCCCCAGTTCTGTTCCCGCGCTACGAGAGCAGCGCCAGATGATGTCGCTG 578
 QY 764 CCGCGGGGGCGCGTGCCTGCGGTGTTGGAAACGTACAGCCGCGGCTGGTGTATGACGG 823
 Db 577 CCGCGGGGGCGCGTGCCTGCGGTGTTGGAAACGTACAGCCGCGGCTGGTGTATGACGG 518
 QY 824 TACGGCGACGGGGCGGCTACTCCCGCGGTGCTGCTGGCGCGGAGAGGGCTGGCGCT 893
 Db 517 TACGGCGACGGGGCGGCTACTCCCGCGGTGCTGCTGGCGCGGAGAGGGCTGGCGCT 458
 QY 884 CTCCTCAGCGGGACGGGGTTCCGTGAGGAGAGCAGCCGCGGGTGAGCGCCGGGCTTC 943
 Db 457 CTCCTCAGCGGGACGGGGTTCCGTGAGGAGAGCAGCCGCGGGTGAGCGCCGGGCTTC 398
 QY 944 CTTGAACCTCTCCAGGACACCGCCCTCCCGCCAGCGTGCCACCCGACCTTCGCGCGGC 1003
 Db 397 CTTGAACCTCTCCAGGACACCGCCCTCCCGCCAGCGTGCCACCCGACCTTCGCGCGGC 338
 QY 1004 GGCATCAGAGCGCTGTGTCACCGTACACGAGAGGGCCCTGAGAGGGCGCCACGCGGC 1063
 Db 337 GGCATCAGAGCGCTGTGTCACCGTACACGAGAGGGCCCTGAGAGGGCGCCACGCGGC 278
 QY 1064 CAGGGCGCCCTCAGGGGTGCTGGAATCTGTCCGACACCCACGAGCAGGTGAGCG 1123
 Db 277 CAGGGCGCCCTCAGGGGTGCTGGAATCTGTCCGACACCCACGAGCAGGTGAGCG 218
 QY 1124 CGAGGATCC 1132
 Db 217 CGAGGATCC 209

RESULT 6
 BE793330
 LOCUS 601588311F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3942507 5',
 DEFINITION mRNA sequence.
 BE793330
 ACCESSION BE793330.1 GI:10214528

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 623)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
REFERENCE Contact: Robert Strausberg, Ph.D.
AUTHORS Email: cgapbs-r@mail.nih.gov
TITLE Tissue Procurement: DCTD/DTF
JOURNAL cDNA Library Preparation: Ling Hong/Rubin Laboratory
COMMENT DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM796 row: e column: 04
 High quality sequence start: 28
 High quality sequence stop: 616.
FEATURES
 Location/Qualifiers
 1..623
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3942507"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_7"
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 37.2%; Score 421.2; DB 2; Length 623;
 Best Local Similarity 99.1%; Pred. No. 6.2e-77;
 Matches 434; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 695 GGGCAGCAGCGTCCCGAGTTCTGCTCCCGCGCTACGAGAGCAGCGCGCAGATGAG 754
 Db 41 GGAGCAGCAGGTCCCGAGTTCTGCTCCCGCGCTACGAGAGCAGCGCGCAGATGAG 100
 QY 755 CTGTCCGTCCCGCGGGGCGCGGTGCGGTGTTGGAAACGTACAGCCGCGCTGGTG 814
 Db 101 CTGTCCGTCCCGCGGGGCGCGGTGCGGTGTTGGAAACGTACAGCCGCGCTGGTG 160
 QY 815 CTATCAGGTACGGGACCGGGGCGCTACTCCCGCGGTGCTGTCGGCGCGGAGGG 874
 Db 161 CTATCAGGTACGGGACCGGGGCGCTACTCCCGCGGTGCTGTCGGCGCGGAGGG 220
 QY 875 CTGGGGCGCTCTCTTACGCGGACCGGGTTCCGTGAGGAGAGACACCCGCGGTGAGGCC 934
 Db 221 CTGGGGCGCTCTCTTACGCGGAGCGGGTTCCGTGAGGAGAGACACCCGCGGTGAGGCC 280
 QY 935 CCGGGGCTTCCCTGAACCTCTCCAGGCCACCGGCCCTTCCCGCCACCGTGTCCACCGACCT 994
 Db 281 CCGGGGCTTCCCTGAACCTCTCCAGGCCACCGGCCCTTCCCGCCACCGTGTCCACCGACCT 340
 QY 995 TCGCGGGGCGCATCCAGAGCGCTGCTGCACCGTACAGCGGCGCTTGGAGCGCGCG 1054
 Db 341 TCGCGGGGCGCATCCAGAGCGCTGCTGCACCGTACAGCGGCGCTTGGAGCGCGCG 400
 QY 1055 CCACGCGCGCAGCGCGCGCTCGAGGGTGGTGGACTCTGTGCGCGCACCCCGACGCGGAG 1114
 Db 401 CCACGCGCGCA-GGCGCGCTCGAGGGTGGTGGACTCTGTGCGCGCACCCCGACGCGGAG 459
 QY 1115 CAGTGAGCGCGAGGATCC 1132

```

Db      460 CAGTGAGCGGAGGATCC 477
|||||
RESULT 7
BG328824
LOCUS      473 bp mRNA linear EST 27-FEB-2001
DEFINITION 602427426F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4546878 5',
            mRNA sequence.
ACCESSION BG328824
VERSION   BG328824.1 GI:13135262
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: NIH Intramural Sequencing Center
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCW1234 row: c column: 07
           High quality sequence stop: 471.
FEATURES
     source
       1..473
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:4546878"
         /tissue_type="adenocarcinoma cell line"
         /lab_host="DH10B (phage-resistant)"
         /clone_lib="NIH_MGC_15"
         /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
         EcoRI; cDNA made by oligo-dT priming. Directionally
         cloned into EcoRI/XhoI sites using the following 5'
         adaptor: GGCACGAG(G). Size-selected >500bp for average
         insert size 1.8kb. Library constructed by Ling Hong in
         the laboratory of Gerald M. Rubin (University of
         California, Berkeley) using ZAP-cDNA synthesis kit
         (Stratagene) and Superscript II RT (Life Technologies)"
ORIGIN
Query Match      37.2%; Score 421; DB 4; Length 473;
Best Local Similarity 97.4%; Pred. No. 6.7e-77;
Matches 449; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 672 GCGGGAGGAGGCGCGCTCCCTAGGAGCAGCGGTCCCGAGTTCTGTGCTTCCCGCGCT 731
Db 1 GCGGGAGGAGGCGCGCTCCCTAGGAGCAGCGGTCCCGAGTTCTGTGCTTCCCGCGCT 60

QY 732 ACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCCGCGGGGCGCGCTGCGCGTGTGG 791
Db 61 ACGAGAGCAGCGCGCAGATGAGCTGTCCGTGCCCGCGGGGCGCGCTGCGCGTGTGG 120

QY 792 AAACGTACAGCGCGCTGTGGCTATGCAGTACGGGACCGGGCGGCGCTACTTCCCG 851
Db 121 AAACGTACAGCGCGCTGTGGCTATGCAGTACGGGACCGGGCGGCGCTACTTCCCG 180

QY 852 CGGTCTCTGCGGCGGAAAGGCTGGCGCTCTCTCTGAGCGGGACGGGTTCCGTGGAG 911
Db 181 CGGTCTCTGCGGCGGAAAGGCTGGCGCTCTCTCTGAGCGGGACGGGTTCCGTGGAG 240

QY 912 GAGACGACCGGCGGGTGAGGCGCGGGGCTTCCCTGAACCTCCACGAGCCACCGCCCTC 971
Db 241 GAGACGACCGGCGGGTGAGGCGCGGGGCTTCCCTGAACCTCCACGAGCCACCGCCCTC 299

```

```

QY 972 CCCCCACCGTGCACCGACCTTCGCGGGGCGCCATCCAGAGCGCGTGTGTCACCGTCA 1031
Db 300 CCCCCACCGTGCACCGACCTTCGCGGGGCGCCATCCAGAGCGCGTGTGTCACCGTCA 359

QY 1032 CACGAGGCGCGCTGAGCGGGGCGCCACGCGCGCGCGCGCGCTCGAGGGTCCGTGGACT 1091
Db 360 CACGAGGCGCGCTGAGCGGGGCGCCACGCGCGCGCGCGCGCTCGAGGGTCCGTGGACT 419

QY 1092 CTGTGCGCGACCCACGAGCGGAGCAGTGTGAGCGCGAGGATCC 1132
Db 420 CTGTGCGCG-ACCCACGAGCGGAGCAGTGTGAGCGCGAGGATCC 459

RESULT 8
AK009605
LOCUS      1571 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
            library, clone:2310034C04 product:2310034C04RIK PROTEIN (SNX28)
            homolog [Mus musculus], full insert sequence.
ACCESSION AK009605
VERSION   AK009605.1 GI:12844500
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS   Carninci, P. and Hayashizaki, Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636

REFERENCE 2
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL   20499374
MEDLINE   11042159
PUBMED    11042159

REFERENCE 3
AUTHORS   Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
            Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multipipillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL   20530913
MEDLINE   11076861
PUBMED    11076861

REFERENCE 4
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            PANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
            Nature 409, 685-690 (2001)
JOURNAL   20530913
MEDLINE   11076861
PUBMED    11076861

REFERENCE 5
AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
JOURNAL   20530913
MEDLINE   11076861
PUBMED    11076861

REFERENCE 6
AUTHORS   Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasekawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
            Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

```



```
/clone lib="NCI CGAP_SG2"
/Note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
```

ORIGIN	
Query Match	33.2%; Score 376.2; DB 4; Length 904;
Best Local Similarity	70.9%; Pred. No. 1.5e-67;
Matches 543; Conservative	0; Mismatches 218; Indels 6; Gaps 3;
QY	13 CCCCCGATACCCAGTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAAGAGGCTCCAAAC 72
Db	1 CCAAAGACCCAGATATCAGCCATGCTGTAGCTTGGTGAAATGGACCGACTCCAGAC 60
QY	73 GTTTCCTCTCTGTCGCTGTGTGAGACGGCAGCAGACCTTCTGTCGCGAGAGTTGGGA 132
Db	61 ATTTCCTCTCTGTCGCTGTGTGAGACCAACAGTGACACATTTGTACGGAGAGCTGGGA 120
QY	133 CGAATTCAGGAGCTCAAGAACCTTCAAGAGACCTTCCGGTGGAGCGGGCTGCT 192
Db	121 TGAGTTCAGGAGCTCCAGAACCTTAAAGAAACCTTCCAGTGGAGCGGCTGCT 180
QY	193 GCGGAGATCTGACCGGCTTCTCCAAAGCTTCTCGATGCACCACTGTTGGGACGCTGGG 252
Db	181 ACGGAGATCTGAACAGTTCTTCCAGCTTCTGATGCTCCATTTGCTGACAGCTCGGGG 240
QY	253 GCGCAGCAGCCCGCGCTGCGCGCTGCGAGCTGTTGGAAACCTATTCTCGGAGGCTGCT 312
Db	241 GCATCTGCTCGAGGACTGTACGTTTGGGCTGCTGGACACCTATGTACAGGCAATTGCT 300
QY	313 GCGGACTGACAGCGGTGGCAGGAGCCCGAGCATCTGCTGCTTCTCGACCCGCAACC 372
Db	301 GGCAACCTCAGAACACATATTGAGGAGTTCAGCACTTTCACGCGCTTCTTGTACCCAAACC 360
QY	373 CTGTGACCTGAGCCCGCGCTGCCACCGGAGCGGGTATCTCTGCCACCCACAGAGGA 432
Db	361 TGTGATCTGGAGCCCATGCTGCTCTGGAGCTGTGATCTGCTGACACAGAGG- 419
QY	433 GCAGCTCTTTCTCGCGCTGCGGGCGCTCTCCATCCACAGTCTGGAGGCTCAGAGCT 492
Db	420 --AGCCCTTATCCCAACCCAGAGCGAGCCTTGACATTCACAGCTGGAGGCTCAGAGCAT 477
QY	493 GGCCTGCTGACGCTCTGTATCCAGGACAGCGGGAATAGGCTTTTCAGGCGCAGGC 552
Db	478 TCCCTGTGTACAGTCTTTCCACACTCTTGACATTAAGAGACAGACACCTTCCACCAAGGC 537
QY	553 CCAGGAGAGCTGGAGCTGCTGCTGCGCACCCCTCAGGCTGGTGGTGGAGAACGA 612
Db	538 TCAGAAATCTGGACATATTAAGCATCTTCAGGCTGGTGGTGGAGAACAA 597
QY	613 AGACGGCAGACCGCTGTTTCCAGCGCTTACCTGGAGGAGCGGCCCG- GGCCAAAG 671
Db	598 GGATCAGCAGGTAGCTGTTTCCAGCTCCCTACCTGGAGGAGGTAGCAACGTCGCCAAG 657
QY	672 GCCGGAGGAGGCGCGCTTCCAGGAGCAGCGGTCCCGAGTTCTGTGCT--TCCCGGCG 729
Db	658 GCCAGAGTTCAGGCTGGCTTTTGCAGAGGAGTGGGAGGAGTCTTGGCAATACCCAAAGGC 717
QY	730 CTACGAGAGCGCCGCGAGATGAGCTGTCCGTGCCCGGGGGCGC 776
Db	718 CTACGAGGCGACGCGCTCTGATGAGCTAATCCGGGGGCTTCAGGGGC 764

```
RESULT 10
BI409553
LOCUS 602963963F1 NCI CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119670 5',
DEFINITION mRNA sequence.
ACCESSION BI409553
VERSION BI409553.1 GI:15170476
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
```

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 726)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11291 row: m column: 15 High quality sequence start: 7 High quality sequence stop: 724. Location/Qualifiers 1..726 /organism="Mus musculus" /mol_type="mRNA" /strain="CZECH II" /db_xref="taxon:10090" /clone="IMAGE:5119670" /tissue_type="pooled lung tumors" /lab_host="DH10B (phage-resistant)" /clone_lib="NCI CGAP Lu33" /note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAATGGAGCGCGCTCTGTTTTTTTTTTTTTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
FEATURES	source
ORIGIN	
Query Match	30.7%; Score 347; DB 4; Length 726;
Best Local Similarity	70.9%; Pred. No. 1.5e-61;
Matches 518; Conservative	0; Mismatches 205; Indels 8; Gaps 4;
QY	12 GCCCCCGATACCCAGTTTTCAGTCAAGGGGCGAGCCCTGGTGCAGATCAAGAGGCTCCAAA 71
Db	3 GGGCCCAAGACACCCAGTATGCCCATGTTGTAGCCTTGTGCA-ATGGACCGACTCCAGA 61
QY	72 CGTTTGCCTTCTGTGTCGCTGGTGCAGCGGAGC-GACACCTTCTGTCGGCAGGAGTTGG 130
Db	62 CATTTGCCCTTCTCGGTGTGCTGGTGCAGACAAACAGGTGACACATTTGTACGGAGGAGCTGG 121
QY	131 GACCAATTCAGCAGCTCAAGAGACCTTCAAGGAGACCTTCCCGTGGAGCGGCGCTG 190
Db	122 GATGAGTTTCAGGAGCTC---CAGACCTTAAAGAAACCTTCCCACTGAGGAGCGGCTG 178
QY	191 CTGGGAGATCTGACCGGCTTCTCCAAAGCTTCTCGATGCACCACTGTTGGAGCGCGTG 250
Db	179 CTACGGAGATCTGAACAAGTTCTTCCCAAGCTTCTGATGCTCCATGCTGACACGTCGG 238
QY	251 GGGCGCACGACCGCGGCTTGGCGGCTGAGCTGTTGGAAACCTATTCTCGGAGGCTG 310
Db	239 GGGCATACTGTCGAGGACTGCTACGTTTGGCGGCTGCTGGACACCTATGTACAGGCATTG 298
QY	311 CTGGCATCTGAGCGCGTGGCAGCGGACCCGACGATCACTGGCTTCTTGCACCGGAA 370
Db	299 CTGGCAACTCAGAACACATATTAGGAGTTTCAGCACTTCACGGCTTCTTTGTATCCCAA 358
QY	371 CCCCTGGACCTGGAGCCCGCGCTGCCACCCGCGAGCGGGTGATCTCTGCCACCCAGAG 430
Db	359 CCTCTGGATCTGGAGCCCATCTGCCTCTGGCAGCCTGGTGATCCTGCCTACACGAGAG 418


```
Db 499 GGCTCAGAGATTCCCTGTGTACAGCCTTTCCACACTCTTTGACATAGAGACAGACCTTT 558
QY 541 TCAGGGCGCAGCCAGAGAGCCTGACGCTGCTGCTGCGGCAC--CCCTCAGGCTGG-TG 597
Db 559 CCACACCAAGGCTCAAGAAATTTGGCACATTTACTTACGCACATCTTTCAGGCTGGTTG 618
QY 598 GCTGTGTGGAGAACGAAGA-CCGCGACAGCCGCTGGTTT-CCAGGCGCCTACCT-GGAGGA 654
Db 619 GCTGTGTGGAGAACGAAGATCACGAGGTAGCTGTGGTTTCCACAGATCCCTACCTGGGAGGA 678
QY 655 GCGCGCCCGGGCCAAAGCCGCGGAGGA-----GGCCGCTCCTAGGGAGACCGGTCC 708
Db 679 GGTAGCAAAAGCTCCAGGCGAGAGTCAAGCCTGGCTTTTTCACAGGGAATCGGCAG 738
QY 709 CCAGTCTGTGCTTCCCGCCCTACGAGAGACGCGCAGATCAGCTGTCCGTGCGCCGC 768
Db 739 GCAGTCTGCACCTACCCAGCATACGACGACGTCGCTCTGTATGGAGTATCCGCGCCCGAG 798
QY 769 GG 770
Db 799 GG 800

RESULT 14
CB113007
LOCUS
DEFINITION
CB113007 602 bp mRNA linear EST 28-JAN-2003
K-E0154911 L6ChoCK0 Homo sapiens cDNA clone L6ChoCK0-8-B11 5',
mRNA sequence.
CB113007
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 602)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 8 row: B column: 11
High quality sequence stop: 602.
Location/Qualifiers
1..602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L6ChoCK0-8-B11"
/sex="M"
/cell_line="Cho-CK"
/lab_host="Top10F"
/clone_lib="L6ChoCK0"
/notes="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
```

```
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

Query Match 28.0%; Score 317.4; DB 6; Length 602;
Best Local Similarity 99.7%; Pred. No. 2e-55;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCATGGCAGGCGCCCGATACCCAGTTCCTTCAGTGCAGGGGAGCCCTGGTGAGATCAA 60
Db 284 AGCATGGCAGGCGCCCGATACCCAGTTCCTTCAGTGCAGGGGAGCCCTGGTGAGATCAA 343
QY 61 GAGGCTCCAAAGCTTTGCTCTCTCTGCGCTGGTTCAGACGGCAGACACCTTCGTGCG 120
Db 344 GAGGCTCCAAAGCTTTGCTCTCTCTGCGCTGGTTCAGACGGCAGACACCTTCGTGCG 403
QY 121 CAGGAGTTGGGAGCAATTCAGGAGCTCAAGAGACCTTCAGGAGACCTTCGCGGTGA 180
Db 404 CAGGAGTTGGGAGCAATTCAGGAGCTCAAGAGACCTTCAGGAGACCTTCGCGGTGA 463
QY 181 GCGGCGCTGCTGCGGAGATCTGACCGCTTCCTCCAAAGCTTCTCGATGACCACTGTT 240
Db 464 GCGGCGCTGCTGCGGAGATCTGACCGCTTCCTCCAAAGCTTCTCGATGACCACTGTT 523
QY 241 GGGACCGCTGGGCGCGCACGAGCCGCGCTGGCGCCCTGCAGCTGTGGAAACCTATTTC 300
Db 524 GGGACCGCTGGGCGCGCACGAGCCGCGCTGGCGCCCTGCAGCTGTGGAAACCTATTTC 583
QY 301 TCGGAGGCTGTGGCGCACT 319
Db 584 TCGGAGGCTGTGGCGCACT 602

RESULT 15
BQ935073
LOCUS
DEFINITION
AGENCOURT 8799923 NC1 CGAP_Co24 Mus musculus cDNA clone
IMAGE:6399072 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 967)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13899 row: j column: 01
High quality sequence stop: 572.
Location/Qualifiers
1..967
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6399072"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life

FEATURES
source

ORIGIN

Search completed: May 30, 2005, 23:16:24
Job time : 5927 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:23:17 ; Search time 162 Seconds
(without alignments)
885.729 Million cell updates/sec

Title: US-10-621-113-4

Perfect score: 1950

Sequence: 1 MAGPRYPVSGAALVQIKR.....RRQRPGRGCVDSVPHFTTEQ 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1950	100.0	371	8 ADJ57071	Adj57071 Nox enzym
2	1937.5	99.4	376	8 ADJ57073	Adj57073 Nox enzym
3	1934.5	99.2	370	8 ADJ57069	Adj57069 Nox enzym
4	1922	98.6	375	8 ADJ57075	Adj57075 Nox enzym
5	1418	72.7	419	4 ABG24891	Abg24891 Novel hum
6	1418	72.7	726	4 ABG25247	Abg25247 Novel hum
7	309.5	15.9	390	8 ADG36779	Adg36779 Human P47
8	307.5	15.8	390	8 ADG36778	Adg36778 Human P47
9	307.5	15.8	390	8 ADG62962	Adg62962 p47 (phox)
10	307.5	15.8	390	8 ADP23795	Adp23795 PRO polyp
11	307	15.7	389	8 ADG36776	Adg36776 Human P47
12	306.5	15.7	390	8 ADL14450	Adl14450 Human NF-
13	306.5	15.7	976	4 ABG20715	Abg20715 Novel hum
14	306	15.7	389	8 ADG36774	Adg36774 Rat P47PH
15	304.5	15.6	366	8 ABM80885	Abm80885 Tumour-as
16	304.5	15.6	390	2 AAR83825	Aar83825 p47 (phox)
17	304.5	15.6	1379	4 ABG20711	Abg20711 Novel hum
18	296	15.2	1619	4 ABG20707	Abg20707 Novel hum
19	278.5	14.3	386	6 AAO16002	Aao16002 Human NAD
20	246.5	12.6	1124	6 AAE38441	Aae38441 Mouse FIS
21	211	10.8	1054	6 ADA55008	Ada55008 Human pro
22	178	9.1	278	8 ABM80892	Abm80892 Tumour-as
23	163	8.4	268	8 ABU17156	Abu17156 Novel sig
24	163	8.4	268	7 ADB93864	Adb93864 Human nov
25	163	8.4	760	6 AAE34861	Aae34861 Human kin

Aae38440 Human FIS
Aau87356 Novel cen
Adi54671 Novel hum
Aau68543 Human nov
Aau14174 Human nov
Abu10058 Human NAD
Abu10044 Human NAD
Aag64031 Human PX
Aag64033 Human PX
Aaw87504 Human N-m
Aaw87503 Human N-m
Adn21559 Bacterial
Aap70196 Sequence
Aay92020 Human mul
Abb82988 Human mul
Adh11597 Human bon
Aab26240 Human N-m
Abb55523 Human NMD
Abu61440 Human N-m
Aae39251 Human NMD

ALIGNMENTS

RESULT 1
ADJ57071
ID ADJ57071 standard; protein; 371 AA.
XX
AC ADJ57071;
XX
DT 06-MAY-2004 (first entry)
XX
DE Nox enzyme regulatory protein p41Nox.
XX
KW Nox regulatory protein; p41Nox; NADH oxidase; cytosolic; human.
XX
OS Homo sapiens.
XX
PN WC2004007689-A2.
XX
PD 22-JAN-2004.
XX
PF 16-JUL-2003; 2003WO-US022246.
XX
PR 16-JUL-2002; 2002US-0396170P.
PR 23-AUG-2002; 2002US-0405647P.
XX
(UYEM-) UNIV EMORY.
XX
Lambeth JD, Cheng G;
XX
WFI; 2004-122927/12.
XX
N-PSDB; ADJ57070.
XX
New regulatory protein for hydrogenated nicotinamide adenine dinucleotide phosphate oxidase, useful for preparing a medicament for affecting superoxide formation or regulating cellular proliferation in an animal or a human.
XX
Claim 1; SEQ ID NO 4; 70pp; English.
XX
The present sequence is the protein sequence of a human regulatory protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and nucleotide sequences of four p41Nox variants have been identified
XX
ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
XX
a PX domain and 2 SH3 domains and show sequence homology (25% identity)
XX
with p47phox, a regulatory protein for gp91phox. They function in the
XX
regulation of cell growth and are therefore implicated in diseases
XX
involving abnormal cell growth, such as cancer. They may also function in
XX
innate immune mechanisms of epithelial tissue or other barrier cells, and
XX
hence may be involved in diseases of diminished ability to fight
XX
infections or inflammatory conditions. p41Nox regulatory proteins and

CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.

XX Sequence 371 AA;

Query Match 100.0%; Score 1950; DB 8; Length 371;
 Best Local Similarity 100.0%; Pred. NO. 4.1e-169;
 Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGDTFVRRSWDEFRQLKTKLKEFPVEA 60
 DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGDTFVRRSWDEFRQLKTKLKEFPVEA 60

QY 61 GLLRRSDRVLPKLLDAPLLGRVGRTSRGLARLQLLETYSRRLATAERVARSPITIGFFA 120
 DB 61 GLLRRSDRVLPKLLDAPLLGRVGRTSRGLARLQLLETYSRRLATAERVARSPITIGFFA 120

QY 121 POPLDEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTRDRPFQ 180
 DB 121 POPLDEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTRDRPFQ 180

QY 181 AQAQESLDVLLRHPSGMWLVENEDQTAWFPAPYLEEAAPGQREGGSLGSSGPGQFCAS 240
 DB 181 AQAQESLDVLLRHPSGMWLVENEDQTAWFPAPYLEEAAPGQREGGSLGSSGPGQFCAS 240

QY 241 RAYESSRADELSPAGARVRVLETSDRGMWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
 DB 241 RAYESSRADELSPAGARVRVLETSDRGMWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300

QY 301 RCGDDPAGEARGFPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPRGRGRC 360
 DB 301 RCGDDPAGEARGFPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPRGRGRC 360

QY 361 VDSVPHPTTEQ 371
 DB 361 VDSVPHPTTEQ 371

RESULT 2
 ADJ57073

ID ADJ57073 standard; protein; 376 AA.

AC ADJ57073;

DT 06-MAY-2004 (first entry)

DE Nox enzyme regulatory protein p41Nox.

DE Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.

KW Homo sapiens.

OS WO2004007689-A2.

PN 22-JAN-2004.

PD 16-JUL-2003; 2003WO-US022246.

PF 16-JUL-2002; 2002US-0396170P.

PR 23-AUG-2002; 2002US-0405647P.

XX (UYEM-) UNIV EMORY.

PI Lambeth JD, Cheng G;

XX WPI; 2004-122927/12.

DR N-PSDB; ADJ57072.

PT New regulatory protein for hydrogenated nicotinamide adenine dinucleotide

PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.

PS Claim 1; SEQ ID NO 6; 70pp; English.

XX The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.

XX Sequence 376 AA;

Query Match 99.4%; Score 1937.5; DB 8; Length 376;
 Best Local Similarity 98.7%; Pred. NO. 5.7e-167;
 Matches 371; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGDTFVRRSWDEFRQLKTKLKEFPVEA 60
 DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGDTFVRRSWDEFRQLKTKLKEFPVEA 60

QY 61 GLLRRSDRVLPKLLDAPLLGRVGRTSRGLARLQLLETYSRRLATAERVARSPIT 115
 DB 61 GLLRRSDRVLPKLLDAPLLGRVGRTSRGLARLQLLETYSRRLATAERVARSPIT 120

QY 116 TGFAPQPLDEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR 175
 DB 121 TGFAPQPLDEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPFCTQDTR 180

QY 176 DRPFOAQAESLDVLLRHPSGMWLVENEDQTAWFPAPYLEEAAPGQREGGSLGSSGP 235
 DB 181 DRPFOAQAESLDVLLRHPSGMWLVENEDQTAWFPAPYLEEAAPGQREGGSLGSSGP 240

QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGMWLCRYGDRAGLLPAVLLRPEGLGALL 295
 DB 241 QFCASRAYESSRADELSPAGARVRVLETSDRGMWLCRYGDRAGLLPAVLLRPEGLGALL 300

QY 296 SGTGFRGGDDPAGEARGFPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPRQG 355
 DB 301 SGTGFRGGDDPAGEARGFPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRRRPRQG 360

QY 356 RPRGCVDSVPHPTTEQ 371
 DB 361 RPRGCVDSVPHPTTEQ 376

RESULT 3

ADJ57069

ID ADJ57069 standard; protein; 370 AA.

AC ADJ57069;

DT 06-MAY-2004 (first entry)

DE Nox enzyme regulatory protein p41Nox.

XX Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.

OS Homo sapiens.

PN WO2004007689-A2.
 XX 22-JAN-2004.
 XX 16-JUL-2003; 2003WO-US022246.
 XX 16-JUL-2002; 2002US-0396170P.
 PR 23-AUG-2002; 2002US-0405647P.
 XX (UYEM-) UNIV EMORY.
 XX Lambeth JD, Cheng G;
 XX WPI; 2004-122927/12.
 DR N-PSDB; ADJ57068.
 XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
 PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.
 XX
 XX Claim 1; SEQ ID NO 2; 70pp; English.
 XX
 CC The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.
 XX
 XX Sequence 370 AA;
 XX
 Query Match 99.2%; Score 1934.5; DB 8; Length 370;
 Best Local Similarity 99.7%; Pred. No. 1e-166;
 Matches 370; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRSWDFRQLKTLKETPVEA 60
 DB 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRSWDFRQLKTLKETPVEA 59
 QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGRVTSRGLARLQLLETYSRLLATAERVARSP 120
 DB 60 GLLRRSDRVLPKLGDASLDAPLLGRVGRVTSRGLARLQLLETYSRLLATAERVARSP 119
 QY 121 PQLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFQ 180
 DB 120 PQLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPFQ 179
 QY 181 AQAQESLDVLLRHPSGWLNVENEDROTAPFPAPYLEAAPGQREGGPGSLGSSGPOFCAS 240
 DB 180 AQAQESLDVLLRHPSGWLNVENEDROTAPFPAPYLEAAPGQREGGPGSLGSSGPOFCAS 239
 QY 241 RAYESSRADELSVPAGARVRLTSDRGWMLCVYGDAGLLPAVLARPEGLGALLSGTG 300
 DB 240 RAYESSRADELSVPAGARVRLTSDRGWMLCVYGDAGLLPAVLARPEGLGALLSGTG 299
 QY 301 RGGDDPAGEARGPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRPRQGRPGC 360
 DB 300 RGGDDPAGEARGPPEPSQATAPPTVTPRSPGAIQSRCTVTRRALERRPRQGRPGC 359
 QY 361 VDSVPHPTTEQ 371
 |||||||

Db 360 VDSVPHPTTEQ 370
 RESULT 4
 ADJ57075
 ID ADJ57075 standard; protein; 375 AA.
 XX AC ADJ57075;
 XX DT 06-MAY-2004 (first entry)
 XX DE Nox enzyme regulatory protein p41Nox.
 XX KW Nox regulatory protein; p41Nox; NADH oxidase; cytostatic; human.
 XX OS Homo sapiens.
 XX PN WO2004007689-A2.
 XX PD 22-JAN-2004.
 XX PF 16-JUL-2003; 2003WO-US022246.
 XX PR 16-JUL-2002; 2002US-0396170P.
 PR 23-AUG-2002; 2002US-0405647P.
 XX PA (UYEM-) UNIV EMORY.
 XX PI Lambeth JD, Cheng G;
 XX WPI; 2004-122927/12.
 DR N-PSDB; ADJ57074.
 XX New regulatory protein for hydrogenated nicotinamide adenine dinucleotide
 PT phosphate oxidase, useful for preparing a medicament for affecting
 PT superoxide formation or regulating cellular proliferation in an animal or
 PT a human.
 XX
 XX Claim 1; SEQ ID NO 8; 70pp; English.
 XX
 CC The present sequence is the protein sequence of a human regulatory
 CC protein, denoted p41Nox, for Nox (NADH oxidase) enzymes. The protein and
 CC nucleotide sequences of four p41Nox variants have been identified
 CC ADJ57068-ADJ57075. p41Nox proteins have a mol.wt. of 41 kDa. They include
 CC a PX domain and 2 SH3 domains and show sequence homology (25% identity)
 CC with p47phox, a regulatory protein for gp91phox. They function in the
 CC regulation of cell growth and are therefore implicated in diseases
 CC involving abnormal cell growth, such as cancer. They may also function in
 CC innate immune mechanisms of epithelial tissue or other barrier cells, and
 CC hence may be involved in diseases of diminished ability to fight
 CC infections or inflammatory conditions. p41Nox regulatory proteins and
 CC nucleic acids are useful for preparing a medicament for affecting
 CC superoxide formation or regulating cellular proliferation in an animal or
 CC a human. The regulatory proteins are also useful in drug development,
 CC e.g. screening for drugs that regulate the biological activity of the
 CC proteins, or in assays that relate to assessment of abnormal growth or
 CC cellular proliferation including cancer.
 XX
 XX Sequence 375 AA;
 XX
 Query Match 98.6%; Score 1922; DB 8; Length 375;
 Best Local Similarity 98.4%; Pred. No. 1.4e-165;
 Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;
 QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRSWDFRQLKTLKETPVEA 60
 DB 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSGSDTFVRSWDFRQLKTLKETPVEA 59
 QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGRVTSRGLARLQLLETYSRLLATAERVARSP 115
 DB 60 GLLRRSDRVLPKLGDASLDAPLLGRVGRVTSRGLARLQLLETYSRLLATAERVARSP 119
 QY 116 TGFFAPQPLDLEPALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
 |||||||

Db 120 TGFFAPQPLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 179
QY 176 DRPFQAQAESLDVLLRHPSGWLWVENEDROTAMFPAPYLEAAPQGREGPSLGSSGP 235
Db 180 DRPFQAQAESLDVLLRHPSGWLWVENEDROTAMFPAPYLEAAPQGREGPSLGSSGP 239
QY 236 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLIPAVLIRPEGIGALL 295
Db 240 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLIPAVLIRPEGIGALL 299
QY 296 SGTGFRGDDPAGEARGFPEPSQATAPPTVTRPSGAIQSRCCCTVTRALLERPRRQG 355
Db 300 SGTGFRGDDPAGEARGFPEPSQATAPPTVTRPSGAIQSRCCCTVTRALLERPRRQG 359
QY 356 RPRGCVDSVPHTTEQ 371
Db 360 RPRGCVDSVPHTTEQ 375

RESULT 5
ABG24891
ID ABG24891 standard; protein; 419 AA.
XX AC ABG24891;

XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24882.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS89078.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 55250; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 419 AA;

QY Query Match 72.7%; Score 1418; DB 4; Length 419;
Best Local Similarity 95.9%; Pred. No. 9.1e-120;
Matches 278; Conservative 0; Mismatches 4; Indels 8; Gaps 2;

QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSTFVRRSWDEPRLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSTFVRRSWDEPRLKTKLKEFPVEA 60

QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGTSGRLARLQLLETYSRRLLTAAERVARSPIT 115

Db 61 GLLRRSDRVLPKLQASLDAPLLGRVGTSGRLARLQLLETYSRRLLTAAERVARSPIT 120

QY 116 TGFFAPQPLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 175

Db 121 TGFFAPQPLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 180

QY 176 DRPFQAQAESLDVLLRHPSGWLWVENEDROTAMFPAPYLEAAPQGREGPSLGSSGP 235

Db 181 DRPFQAQAESLDVLLRHPSGWLWVENEDROTAMFPAPYLEAAPQGREGPSLGSSGP 240

QY 236 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRYGDRAGLIPAVL 285

Db 241 QFCASRAYESSRADELSVPAGARVRVLETSRGMWLCRY---AGAGPEEL 287

RESULT 6

ABG25247

ID ABG25247 standard; protein; 726 AA.

XX AC ABG25247;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #25238.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS89434.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 55606; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIFO at ftp.wifo.int/pub/published_pct_sequences

XX SQ Sequence 726 AA;

Query Match 72.7%; Score 1418; DB 4; Length 726;
Best Local Similarity 95.9%; Pred. No. 1.9e-119;
Matches 278; Conservative 0; Mismatches 4; Indels 8; Gaps 2;

QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDFPQLKTKLTPVEA 60
DB 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDFPQLKTKLTPVEA 60

QY 61 GLLRSDRVLPKL-----LDAPLLGRVGRTRGLARLQLLETYSRRLATAERVARSP 115
DB 61 GLLRSDRVLPKLLGQASLDAPLLGRVGRTRGLARLQLLETYSRRLATAERVARSP 120

QY 116 TGFFAPQPLDLEPALPPGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 175
DB 121 TGFFAPQPLDLEPALPPGSRVILPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTR 180

QY 176 DRPFOAQAESLDVLRHPSGWLNVENEDROTAWFPAPYLEEAAPQGGREGGSLGSSGP 235
DB 181 DRPFOAQAESLDVLRHPSGWLNVENEDROTAWFPAPYLEEAAPQGGREGGSLGSSGP 240

QY 236 QFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVL 285
DB 241 QFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRY---AGAGPEEL 287

RESULT 7
ADG36779
ID ADG36779 standard; protein; 390 AA.
XX
AC ADG36779;
XX
XX 26-FEB-2004 (first entry)
XX
DE Human P47PHOX-related protein.
XX
KW human; autoimmune condition; NADPH oxidase activity;
KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
KW cardiovascular disease.
XX
OS Homo sapiens.
XX
XX WO2003095667-A2.
XX
XX 20-NOV-2003.
XX
XX 13-MAY-2003; 2003WO-IB002419.
XX
XX 13-MAY-2002; 2002US-0380904P.
PR 27-NOV-2002; 2002US-0429609P.

XX (AREX-) AREXIS AB.
XX Holmdahl R, Olofsson P;
XX WPI; 2004-012133/01.
DR
XX Assessing a mammal's susceptibility to develop an autoimmune condition by
PT determining whether or not the level of NADPH oxidase activity of the
PT cell after contacting the cell with an NADPH oxidase activator is less
PT than a control level.
XX
PS Disclosure; SEQ ID NO 7; 103pp; English.
XX
XX The invention comprises a method for assessing the susceptibility of a
CC mammal to develop an autoimmune condition. The method involves: providing
CC a blood or synovial fluid sample containing a cell from a mammal;
CC determining the level of NADPH oxidase activity of the cell after
CC contacting the cell with an NADPH oxidase activator, and determining
CC whether or not the level is less than a control level of NADPH oxidase
CC activity. The method further comprises determining whether or not a
CC mammal has a genetic variant of the gene encoding a polypeptide that
CC functions in the NADPH oxidase pathway (e.g. P47PHOX), where the presence
CC of the genetic variant indicates that the mammal is susceptible to
CC develop an autoimmune condition. The method is useful in the diagnosis
CC and treatment of autoimmune conditions, such as: arthritis, multiple
CC sclerosis, lupus, autoimmune uveitis, type I diabetes, bronchial asthma,
CC septic arthritis and cardiovascular disease. The present amino acid
CC sequence represents a human P47PHOX-related protein.

XX SQ Sequence 390 AA;

Query Match 15.9%; Score 309.5; DB 8; Length 390;
Best Local Similarity 25.5%; Pred. No. 3.2e-19;
Matches 96; Conservative 62; Mismatches 151; Indels 67; Gaps 13;

QY 24 FAFSVRWSDGSDTFVRRSWDFPQLKTKLTFPVEAGLLRSDRVLPKLLDAPLLGRVG 83
DB YMFVLKWDLSKVVVRRFTETVEFHKTLKEMFPIEAGAINPENRIIPH-LPAKWFDDG 84

QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITTFAPQPLDLEPALPPGSRV-----I 137
DB RAAEN--RQGTLETCSTLMSLPTKISRCPHLLDFKVPDDLK--LPTDNQTKKPETYL 140

QY 138 LPTPEOPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTRDRPFOAQAESLDVLRHPSGW 197
DB MPDQKSTATDITGPILL-----QSYRAIANY--EKTSGSENALSTGCVVEVVEKSESW 193

QY 198 WLVENEDROTAWFPAPYLEEA-APQGGREGGSLGSSGPQFCASRAYESSRADELSVPAG 256
DB WFCQMKAKR-GWIPASFLEPLDSPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250

QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
DB EAVEVIHKLLDGMVIRKDDVTGYFFSMYLOKSGQDVSOAQRIKRGAPRRSIRNVHS 310

QY 294 -----LLSGTGFRRGGD-----DPAGEARGEP-----EPQATAPPTVTPRS 331
DB IHQRGRKLSDQAYRRNSVRFLOQRRRQARPGFPGSPLEBERQTRQSKPQPAVPPRS 370

QY 332 PGAIQSRCTCTVRRAL 347

DB ADLILNRCSESTKRL 386

RESULT 8
ADG36778
ID ADG36778 standard; protein; 390 AA.
XX
XX ADG36778;
AC
XX
XX 26-FEB-2004 (first entry)
XX

Db 141 MPKDGKSTADITGPIL-----QTVRAIADY--EKTSGSEMALSTGDDVVVVEKSESGW 193
QY 198 WLVENEDROTAWFPAPYLEEA-APGQREGGPGSLGSGPOFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASGLEPLDSPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
QY 257 ARVRVLETSRQGWLCRYGDRAGLLPAVLRLRPG-----LGA----- 293
Db 251 EAVEVIHKLDDGWWVIRKDDVTGYFPMYLOKSGQDVSAQORQIKRGAPPRSSIRNAHS 310
QY 294 -----LLSGTGFRGDD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
Db 311 IHQSRKRLSQDAYRNNSVRFLOQRRRQARPGQSPGSPLEERQTRQSKPQPAVPPRPS 370
QY 332 PGAIQSRCCCTVTRRAL 347
Db 371 ADLILNRCSESTKRKL 386

RESULT 10
ID ADP23795
XX ADP23795 standard; protein; 390 AA.
AC ADP23795;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide SEQ ID NO:973.
DE
KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
OS Unidentified.
XX
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX 30-OCT-2003; 2003WO-US034312.
XX
XX 01-NOV-2002; 2002US-0423394P.
XX
XX (GETH) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
XX Wu TD;
XX
XX WPI; 2004-419628/39.
XX
XX N-PSDB; ADP23794.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.
XX
XX Claim 7; SEQ ID NO 973; 2940pp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,

CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
XX of the invention.

XX Sequence 390 AA;

Query Match 15.8%; Score 307.5; DB 8; Length 390;
Best Local Similarity 25.3%; Pred. No. 4.9e-19;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

QY 24 FAFSVRWSDGSDTFVRRSWDEFQRLKTKLTFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
Db 26 YMFVLKQDLSEKVVVRRFTEIVEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
QY 84 RTSRGLARQLLETYSRRLATAERVARSPPTTGTFFAPQPLDLEPALPGSRV-----I 137
Db 85 RAAEN--RQGTLTETCYCSTLMSLPTKISRCPHLLDFKVRPDDLK--LPTDNOTKKPETYL 140
QY 138 LPTPEEQPLSRAAGRLSIHLEAQSLRCLQPFCTQDTRDRPFOAQAESLDVLLRHPSGW 197
Db 141 MPKDGKSTADITGPIL-----QTVRAIADY--EKTSGSEMALSTGDDVVVVEKSESGW 193
QY 198 WLVENEDROTAWFPAPYLEEA-APGQREGGPGSLGSGPOFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASGLEPLDSPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
QY 257 ARVRVLETSRQGWLCRYGDRAGLLPAVLRLRPG-----LGA----- 293
Db 251 EAVEVIHKLDDGWWVIRKDDVTGYFPMYLOKSGQDVSAQORQIKRGAPPRSSIRNAHS 310
QY 294 -----LLSGTGFRGDD-----DPAGEARGFP-----EPSQATAPPTVTPRPS 331
Db 311 IHQSRKRLSQDAYRNNSVRFLOQRRRQARPGQSPGSPLEERQTRQSKPQPAVPPRPS 370
QY 332 PGAIQSRCCCTVTRRAL 347
Db 371 ADLILNRCSESTKRKL 386

RESULT 11

ADG36776
ID ADG36776 standard; protein; 389 AA.

AC ADG36776;

DT 26-FEB-2004 (first entry)

XX Human P47PHOX protein #1.

XX human; autoimmune condition; NADPH oxidase activity;

KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
KW cardiovascular disease.

OS Homo sapiens.

XX WO2003095667-A2.

PN 20-NOV-2003.

XX 13-MAY-2003; 2003WO-IB002419.

XX 13-MAY-2002; 2002US-0380904P.

PR 27-NOV-2002; 2002US-0429609P.

XX (AREX-) AREXIS AB.
 XX Holmdahl R, Olofsson P;
 PI WPI; 2004-012133/01.
 DR N-PSDB; ADG36775.
 XX Assessing a mammal's susceptibility to develop an autoimmune condition by
 PT determining whether or not the level of NADPH oxidase activity of the
 PT cell after contacting the cell with an NADPH oxidase activator is less
 PT than a control level.
 XX Claim 18; SEQ ID NO 4; 103pp; English.
 PS
 XX The invention comprises a method for assessing the susceptibility of a
 CC mammal to develop an autoimmune condition. The method involves: providing
 CC a blood or synovial fluid sample containing a cell from a mammal,
 CC determining the level of NADPH oxidase activity of the cell after
 CC contacting the cell with an NADPH oxidase activator, and determining
 CC whether or not the level is less than a control level of NADPH oxidase
 CC activity. The method further comprises determining whether or not a
 CC mammal has a genetic variant of the gene encoding a polypeptide that
 CC functions in the NADPH oxidase pathway (e.g. P47PHOX), where the presence
 CC of the genetic variant indicates that the mammal is susceptible to
 CC develop an autoimmune condition. The method is useful in the diagnosis
 CC and treatment of autoimmune conditions, such as: arthritis, multiple
 CC sclerosis, lupus, autoimmune uveitis, type I diabetes, bronchial asthma,
 CC septic arthritis and cardiovascular disease. The present amino acid
 CC sequence represents a human P47PHOX protein.
 XX
 XX Sequence 389 AA;
 SQ
 Query Match 15.7%; Score 307; DB 8; Length 389;
 Best Local Similarity 27.1%; Pred. No. 5.4e-19;
 Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;
 QY 24 FAFSVRWSDGSTFVRSWDEPRLKKTLETFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
 Db 26 YNFLVKQDLSEKVVYRKFTIYERPHKMLKMFPIERGEIHTENKVIH-LPAPRWYDQ 84
 QY 84 RTSGLARLQLLETYSRELLATAERVARSPITTFGFAPQLDLEPAPGSRVILPTPEE 143
 Db 85 RAAE--SRQGITLTFVNSLMGLPVKISRCPHLLNFKVRPDDLK--LNDQVKKP---E 137
 QY 144 QPLSRAAGRLSIHSLAE---QSLRCLQFCTQDTRDRPFOQAQESLDVLLRHPSGWNL 199
 Db 138 TYLTAKDGNVADITGPILLQTYRAIADY--EKGSKTEMTVATGDVVDVVEKESGWWF 195
 QY 200 VENEDROTAWPAPYLEEA-APQGREGPSLGGSGPQFCASRAVESRDELSPAGAR 258
 Db 196 COMTKR-GWVPASYLEPLDSDPEADDPDNY--AGEPYVTIKAAVVEDEVSLSEGEA 252
 QY 259 VRVLETSRQGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
 Db 253 LEVIHKLGDWVVRKGIITGVFSPMYLQAGEIITQQRIRSGAPRRSTIRNAQSI 312
 QY 294 -----LLSGTGFRGD-----DPAGEARGFPE-----PSQATA-PPPTVTRPSPGA 334
 Db 313 HQSRKRLSQDTRYRNSVRFLQORRRPARPGQSPDSKDNFSTPRAKQPAVPPRPSSDL 372
 QY 335 IQSRCTVTRAL 347
 Db 373 ILHRCSTTKRKL 385
 RESULT 12
 ADRI4450
 ID ADRI4450 standard; protein; 390 AA.
 XX
 AC ADRI4450;
 XX
 DT 21-OCT-2004 (first entry)

XX Human NF-kappaB pathway-associated protein SeqID451.
 DE
 XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnary; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; human.
 XX
 OS Homo sapiens.
 XX WO2004065577-A2.
 PN
 XX 05-AUG-2004.
 PD
 XX 13-JAN-2004; 2004WO-US000798.
 PF
 XX 14-JAN-2003; 2003US-0440068P.
 PR
 XX 12-MAY-2003; 2003US-0469757P.
 PR
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;
 PI
 XX WPI; 2004-562168/54.
 DR N-PSDB; ADRI4451.
 DR
 XX New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 XX Claim 1; SEQ ID NO 451; 237pp; English.
 PS
 XX This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human protein which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

SQ Sequence 390 AA;

Query Match 15.7%; Score 306.5; DB 8; Length 390;
 Best Local Similarity 25.3%; Pred. No. 6e-19;
 Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

QY 24 FAFSVWSDGSDTFVRSWDEFRLQKTKLTFVVEAGLRRSDRVLPKLLDAPLLGRVG 83
 DB 26 YMFVKWQDLSEKVVYRRFTIYEFHKTLEMPFIEAGINPENRIIPH-LPAPKWFDDG 84
 QY 84 RTSRGLARLQLETYSRRLLATAERVARSPITIGFFAPQPLDLEPALPPGSRV-----I 137
 DB 85 RAAEN--RQGLTEYCSLMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKETYL 140
 QY 138 LPTPEQPLSRAAGRLSIHLSLEAQSRLCLQPFCTQDTRDRPFOAQESLDVLLRHPSGW 197
 DB 141 MPKDGKSTATDITGPIL-----QTYRAINY--EKTSGSEMALSTGDDVVVEKSESGW 193
 QY 198 WLVENEDRQTAWFPAPYLEA-APGQREGGSLGSGPOFCASRAYESSRADELSVPAG 256
 DB 194 WFCQMKAKR-GWIPASFLPLSDPDETPENY--AGEPYVAIKAYTAVEGDEVSLLEG 250
 QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
 DB 251 EAVEVIHKLLDGNWVIRKDDVTGYFPSMYLQKSGQDVSOAQRIKGAAPRRSSIRNAHS 310
 QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPPTVTPRPS 331
 DB 311 IHORSRKLSQDAYRRNSVRFLOQRRRQARPGQSPGSPLEERQTRSKPQPAVPPRPS 370
 QY 332 PGAIQSRCCCTVTRRAL 347
 DB 371 ADLILNRCSSESTKRKL 386

RESULT 13
 ABG20715
 ID ABG20715 standard; protein; 976 AA.
 XX AC ABG20715;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #20706.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX XX (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS84902.
 XX XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID NO 51074; 103pp; English.
 XX XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological actions. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences. The invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 976 AA;

Query Match 15.7%; Score 306.5; DB 4; Length 976;
 Best Local Similarity 25.3%; Pred. No. 2e-18;
 Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

QY 24 FAFSVWSDGSDTFVRSWDEFRLQKTKLTFVVEAGLRRSDRVLPKLLDAPLLGRVG 83
 DB 612 YMFVKWQDLSEKVVYRRFTIYEFHKTLEMPFIEAGINPENRIIPH-LPAPKWFDDG 670
 QY 84 RTSRGLARLQLETYSRRLLATAERVARSPITIGFFAPQPLDLEPALPPGSRV-----I 137
 DB 671 RAAEN--RQGLTEYCSLMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKETYL 726
 QY 138 LPTPEQPLSRAAGRLSIHLSLEAQSRLCLQPFCTQDTRDRPFOAQESLDVLLRHPSGW 197
 DB 727 MPKDGKSTATDITGPIL-----QTYRAINY--EKTSGSEMALSTGDDVVVEKSESGW 779
 QY 198 WLVENEDRQTAWFPAPYLEA-APGQREGGSLGSGPOFCASRAYESSRADELSVPAG 256
 DB 780 WFCQMKAKR-GWIPASFLPLSDPDETPENY--AGEPYVAIKAYTAVEGDEVSLLEG 836
 QY 257 ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
 DB 837 EAVEVIHKLLDGNWVIRKDDVTGYFPSMYLQKSGQDVSOAQRIKGAAPRRSSIRNAHS 896
 QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPPTVTPRPS 331
 DB 897 IHORSRKLSQDAYRRNSVRFLOQRRRQARPGQSPGSPLEERQTRSKPQPAVPPRPS 956
 QY 332 PGAIQSRCCCTVTRRAL 347
 DB 957 ADLILNRCSSESTKRKL 972

RESULT 14
 ADG36774
 ID ADG36774 standard; protein; 389 AA.
 XX AC ADG36774;
 XX DT 26-FEB-2004 (first entry)
 XX DE Rat P47PHOX protein.
 XX KW rat; autoimmune condition; NADPH oxidase activity;
 KW NADPH oxidase activator; P47PHOX; arthritis; multiple sclerosis; lupus;
 KW autoimmune uveitis; type I diabetes; bronchial asthma; septic arthritis;
 KW cardiovascular disease.
 XX OS Rattus norvegicus.

Db	2	YMFVWKQDLSEKVVRRFTEIYEFHKTLEKMPFIEAGAINPENRIIPH-LPAPKWFQDQ	60
Qy	81	RVGRTSRGLARQLLETYSRLLATAERVARSPITITGFFAPQPLDLEPALPPGSRV----	136
Db	61	RAAENHOG-----TLTEYCGTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPE	113
Qy	137	--ILPTPEOPLSRAAGRLSIHSLAQSLRCLOPCTQDTRDRPFOAQOESLDVLLRHP	194
Db	114	TYLMPKDGKSTADITGPILL-----QTYRAIANI--EKTSGSEMALSTGDVVVEVEKSE	166
Qy	195	SGHWLVENEDROTAWFPAPVLEEA-APGQREGGPSLGSSGPQFCASRAYESSRADELSV	253
Db	167	SGWFFCOMKAKR-GWIPASFLPLDGPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSL	223
Qy	254	PAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA-----	293
Db	224	LEGEAEVVIHKLDDGWWVIRKDDVTGYFSPMYLQKSGODVSOAQORQIKRGAPPRSSIRN	283
Qy	294	-----LLSGTGFRGDD-----DPAGEARGFP-----EPSQATAPPPTVPT	328
Db	284	VHSIHQSRKRLSQDAYRRNSVRFLQORRRQARFGPQSGPSPLEEERQTQRSKPQPAVPP	343
Qy	329	RSPGAIQSRCCCTVTRAL	347
Db	344	RPSADLILNRCSESTKRKL	362

Search completed: May 28, 2005, 07:40:35
Job time : 165 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:35:33 ; Search time 42 Seconds
(without alignments)
659.400 Million cell updates/sec

Title: US-10-621-113-4

Perfect score: 1950

Sequence: 1 MAGPRYFVSQGAALVQIKR.....RRQRPRGCVDSVPHPTTQ 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307.5	15.8	390	4	US-09-820-005-4
2	307.5	15.8	390	4	US-10-109-856-4
3	278.5	14.3	386	4	US-09-820-005-2
4	278.5	14.3	386	4	US-10-109-856-2
5	153.5	7.9	215	4	US-09-808-701A-21
6	139.5	7.2	560	1	US-07-683-957B-1
7	134	6.9	381	4	US-09-949-016-9555
8	131.5	6.7	614	4	US-09-252-991A-28630
9	130	6.7	532	4	US-09-949-016-7663
10	128.5	6.6	201	4	US-09-949-016-11140
11	128.5	6.6	377	4	US-09-538-092-939
12	128.5	6.6	377	4	US-09-949-016-6693
13	128	6.6	561	4	US-09-252-991A-31113
14	125.5	6.4	1706	4	US-09-252-991A-31760
15	125	6.4	495	4	US-09-252-991A-16992
16	124.5	6.4	462	3	US-08-630-915A-38
17	124.5	6.4	462	4	US-09-879-957-38
18	124.5	6.4	472	4	US-09-252-991A-16723
19	124.5	6.4	520	4	US-09-538-092-1347
20	124	6.4	248	3	US-08-630-915A-40
21	124	6.4	248	4	US-09-879-957-40
22	124	6.4	509	3	US-08-630-915A-194
23	124	6.4	509	4	US-09-879-957-194
24	124	6.4	1676	4	US-09-949-016-7610
25	123.5	6.3	907	4	US-09-252-991A-24114
26	123	6.3	679	4	US-09-252-991A-18857
27	121.5	6.2	917	4	US-09-252-991A-25101

28	119.5	6.1	427	4	US-09-252-991A-30434	Sequence 30434, A
29	119	6.1	554	4	US-09-252-991A-28232	Sequence 28232, A
30	119	6.1	636	4	US-09-252-991A-24902	Sequence 24902, A
31	118.5	6.1	366	4	US-09-252-991A-24338	Sequence 24338, A
32	118.5	6.1	447	3	US-09-199-637A-351	Sequence 351, App
33	118	6.1	847	4	US-09-949-016-6222	Sequence 6222, App
34	118	6.1	885	4	US-09-949-016-7789	Sequence 7789, App
35	118	6.1	1427	4	US-09-252-991A-27005	Sequence 27005, A
36	117.5	6.0	461	4	US-09-252-991A-24717	Sequence 24717, A
37	117	6.0	355	3	US-08-630-915A-192	Sequence 192, App
38	117	6.0	355	4	US-09-879-957-192	Sequence 192, App
39	117	6.0	950	4	US-09-252-991A-25927	Sequence 25927, A
40	116.5	6.0	425	4	US-09-252-991A-17013	Sequence 17013, A
41	115	5.9	526	4	US-09-252-991A-28761	Sequence 28761, A
42	115	5.9	536	4	US-09-252-991A-23495	Sequence 23495, A
43	114.5	5.9	303	4	US-09-538-092-1133	Sequence 1133, App
44	114.5	5.9	486	2	US-08-942-423-2	Sequence 2, Appli
45	114.5	5.9	486	3	US-08-630-915A-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1

US-09-820-005-4

; Sequence 4, Application US/09820005

; Patent No. 6489149

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al

; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; FILE REFERENCE: CLO01198

; CURRENT APPLICATION NUMBER: US/09/820,005

; CURRENT FILING DATE: 2001-03-29

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Human

US-09-820-005-4

Query Match 15.8%; Score 307.5; DB 4; Length 390;
Best Local Similarity 25.3%; Pred. No. 7.5e-21;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

QY	24	PAFSVWSDGDTFVRSWDEFRQLKTKLKEFPVVEAGLLRRSDRVLPKLLDAPLLGRVG	83
DB	26	YMFVKWQDLSEKVVYRFTIYEFHKTLEKMPPIBAGAINPENRIIPH-LPAKWFQDG	84
QY	84	RTSRGLARLQLLETYSRRLLATAERVARSPPTITGRFAPQPLDLEPALPPGSRV-----I	137
DB	85	RAEN--RGTLTEYCVSTLMSLPTKISRCPLHLDFFKVRPDDLK--LPTDNQTKKPYTL	140
QY	138	LPTPEQPLSRAGRLSIHLEAQLRCLOPFCOTDRDPFOAQAESLDVLLRHPSGW	197
DB	141	MPKDGKSTATDITGPIL-----QTVRAIDY--EXTSGSEMAISLTDGVVEVKESEGW	193
QY	198	WLVENEDROTAFWPAPYLEEA-APQGREGGSLGSSGQFCASRAYESSRADELSVPAG	256
DB	194	WFCQMKAKR-GWIPASFLPSDPEDEPEPNY--AGEPYVAIKAYTAVEGDEVSLLEG	250
QY	257	ARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA-----	293
DB	251	EAVEVTHKLDGMMWIRKDDVTGYFPFMYLQSKSQDVQAQRQIKKGAAPRRSIRNAHS	310
QY	294	-----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPTVTPTRPS	331
DB	311	IHQSRKRLSQDAYRENSVRFLOQRROARPPQSPGSLPEEBEQTRQSKPQPAVPRPS	370
QY	332	PGAIOSRCTVTTRAL 347	

```
Db 371 ADLILNRCSSTKRKL 386
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Human
US-10-820-005-2

Query Match 14.3%; Score 278.5; DB 4; Length 386;
Best Local Similarity 24.7%; Pred. No. 4.3e-18;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSDGSDTFVRRSWDEFRLQKTKLTFPPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 YMFVLKQDLSKVVYRFTIYEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 RTSGLARLQLLETYSRRLATAERVARSPITIGFAPQPLDLEPALPPGSRV-----I 137
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 RAAEN--RQGLTEYCSLMSLPTKISRCPHLLDFKVRPDDLK--LPTDNQTKKPEYVL 140
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 LPTPEEQPLSRAAGRLSHLSLEAQSRLCLQPFCTQDTRDRPFOAQAESLDVLLRHPGSW 197
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 MPKDGKSTATDITGPILL-----QTYRAIANY--EKTSGSEMALSTGDVVEVVEKSESGW 193
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 WLVENEDRTAWFPAPYLEEA-APCGREGGPGSLGSSGPOFCASRAYESSRADELSPVAG 256
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 194 WFCQMKAKR-GWIPASFLEPLDSPETEDPEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 EAVEVIHKLLDGMW-----KDDVTGYFFSMYLOKSGQDVSAQORQIKRGAPRRSRSTRNAHS 306
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPPTVTPRPS 331
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 IHQSRKLSQDAYRNRVFLQORRRQARPGQSPGSGPLEEERQTRSKPQPAVPPRPS 366
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 PGAIQSRCTVTRRAL 347
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 367 ADLILNRCSSTKRKL 382

RESULT 4
US-10-109-856-2
; Sequence 2, Application US/10109856
; Patent No. 6709850
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001198DIV
; CURRENT APPLICATION NUMBER: US/10/109,856
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-109-856-2

Query Match 14.3%; Score 278.5; DB 4; Length 386;
Best Local Similarity 24.7%; Pred. No. 4.3e-18;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSDGSDTFVRRSWDEFRLQKTKLTFPPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 YMFVLKQDLSKVVYRFTIYEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 RTSGLARLQLLETYSRRLATAERVARSPITIGFAPQPLDLEPALPPGSRV-----I 137
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 RAAEN--RQGLTEYCSLMSLPTKISRCPHLLDFKVRPDDLK--LPTDNQTKKPEYVL 140
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 LPTPEEQPLSRAAGRLSHLSLEAQSRLCLQPFCTQDTRDRPFOAQAESLDVLLRHPGSW 197
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 MPKDGKSTATDITGPILL-----QTYRAIADY--EKTSGSEMALSTGDVVEVVEKSESGW 193
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 WLVENEDRTAWFPAPYLEEA-APCGREGGPGSLGSSGPOFCASRAYESSRADELSPVAG 256
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 194 WFCQMKAKR-GWIPASFLEPLDSPETEDPEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 EAVEVIHKLLDGMWVIRKDDVTGYFFSMYLOKSGQDVSAQORQIKRGAPRRSRSTRNAHS 310
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 -----LLSGTGFRGGD-----DPAGEARGFP-----EPSQATAPPPTVTPRPS 331
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 311 IHQSRKLSQDAYRNRVFLQORRRQARPGQSPGSGPLEEERQTRSKPQPAVPPRPS 370
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 PGAIQSRCTVTRRAL 347
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 371 ADLILNRCSSTKRKL 386

RESULT 3
US-09-820-005-2
; Sequence 2, Application US/09820005
; Patent No. 6489149
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001198
; CURRENT APPLICATION NUMBER: US/09/820,005
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```


Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11140
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11140

Query Match 6.6%; Score 128.5; DB 4; Length 201;
Best Local Similarity 24.6%; Pred. No. 0.00032;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEEDROTAWFPAPYLE-----E 217
Db 14 YVAQOQELDIKKNERLWLLDDSKSWVRVNSMKNKTFVPSNYVERKNSARKASIVKNLK 73

QY 218 AAPGQGR-EGGPSLSSGPGQFCASRA-----YESSRADELSV 253
Db 74 DTLGIGKVKRKPSPVDS-----ASPADDSFVDPGERLYDLNMPAYVKFNMAEREDELSL 128

QY 254 PACARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGDDPAGEARG 312
Db 129 IKGTKVIVMEKCDGWMRGSYNGQVGWFPNSYVTEE-----GDSPLGDHVG 174

RESULT 11
US-09-538-092-939
; Sequence 939, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 939
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P16333
US-09-538-092-939

Query Match 6.6%; Score 128.5; DB 4; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.00077;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEEDROTAWFPAPYLE-----E 217
Db 13 YVAQOQELDIKKNERLWLLDDSKSWVRVNSMKNKTFVPSNYVERKNSARKASIVKNLK 72

QY 218 AAPGQGR-EGGPSLSSGPGQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVKRKPSPVDS-----ASPADDSFVDPGERLYDLNMPAYVKFNMAEREDELSL 127

QY 254 PACARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGDDPAGEARG 312
Db 128 IKGTKVIVMEKCDGWMRGSYNGQVGWFPNSYVTEE-----GDSPLGDHVG 173

RESULT 12
US-09-949-016-6693
; Sequence 6693, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6693
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6693

Query Match 6.6%; Score 128.5; DB 4; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.00077;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;

QY 179 FOAQOESLDV-----LLRHPGWLWVENEEDROTAWFPAPYLE-----E 217
Db 13 YVAQOQELDIKKNERLWLLDDSKSWVRVNSMKNKTFVPSNYVERKNSARKASIVKNLK 72

QY 218 AAPGQGR-EGGPSLSSGPGQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVKRKPSPVDS-----ASPADDSFVDPGERLYDLNMPAYVKFNMAEREDELSL 127

QY 254 PACARVRVLETSDRGWLWLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGDDPAGEARG 312
Db 128 IKGTKVIVMEKCDGWMRGSYNGQVGWFPNSYVTEE-----GDSPLGDHVG 173

RESULT 13
US-09-252-991A-31113
; Sequence 31113, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31113
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31113

Search completed: May 28, 2005, 07:45:02
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:41:29 ; Search time 141 Seconds
(without alignments)
907.693 Million cell updates/sec

Title: US-10-621-113-4
Perfect score: 1950
Sequence: 1 MAGPRYPVSQGAALVQIKR.....RRQGRPGCVSDVPHPTTEQ 371

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1462099 seqs, 344972447 residues

Total number of hits satisfying chosen parameters: 1462099

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1950	100.0	371	15 US-10-621-113-4	Sequence 4, Appli
2	1937.5	99.4	376	15 US-10-621-113-6	Sequence 6, Appli
3	1934.5	99.2	370	15 US-10-621-113-2	Sequence 2, Appli
4	1922	98.6	375	15 US-10-621-113-8	Sequence 8, Appli
5	309.5	15.9	390	15 US-10-437-427-7	Sequence 7, Appli
6	307.5	15.8	390	14 US-10-109-856-4	Sequence 4, Appli
7	307.5	15.8	390	15 US-10-418-036-18	Sequence 18, Appli
8	307.5	15.8	390	15 US-10-437-427-6	Sequence 6, Appli
9	307.5	15.8	390	16 US-10-767-341-4	Sequence 4, Appli
10	307	15.7	389	15 US-10-437-427-4	Sequence 4, Appli
11	306.5	15.7	390	16 US-10-755-889-451	Sequence 451, App
12	306	15.7	389	15 US-10-437-427-2	Sequence 2, Appli
13	278.5	14.3	386	14 US-10-109-856-2	Sequence 2, Appli

14	278.5	14.3	386	16	US-10-767-341-2	Sequence 2, Appli
15	245.5	12.6	1138	15	US-10-161-927-60	Sequence 60, Appli
16	211	10.8	1054	15	US-10-094-749-2576	Sequence 2576, Ap
17	163	8.4	268	9	US-09-764-868-721	Sequence 721, App
18	158	8.1	204	11	US-09-764-875-874	Sequence 874, App
19	153.5	7.9	215	9	US-09-808-701A-21	Sequence 21, Appli
20	153.5	7.9	215	14	US-10-233-131-21	Sequence 21, Appli
21	153.5	7.9	215	15	US-10-240-145-73	Sequence 73, Appli
22	146	7.5	968	15	US-10-291-265-281	Sequence 281, App
23	143	7.3	134	14	US-10-202-724-4	Sequence 4, Appli
24	141.5	7.3	769	15	US-10-369-493-4212	Sequence 4212, Ap
25	139.5	7.2	560	14	US-10-162-223-7	Sequence 7, Appli
26	139.5	7.2	560	15	US-10-366-345-53	Sequence 53, Appli
27	139.5	7.2	561	9	US-09-813-398-23	Sequence 23, Appli
28	139	7.1	370	10	US-09-946-374-315	Sequence 315, App
29	139	7.1	370	13	US-10-052-596-350	Sequence 350, App
30	139	7.1	370	14	US-10-174-590-350	Sequence 350, App
31	139	7.1	370	14	US-10-176-758-350	Sequence 350, App
32	139	7.1	370	14	US-10-175-737-350	Sequence 350, App
33	139	7.1	370	14	US-10-174-581-350	Sequence 350, App
34	139	7.1	370	14	US-10-176-483-350	Sequence 350, App
35	139	7.1	370	14	US-10-176-749-350	Sequence 350, App
36	139	7.1	370	14	US-10-176-914-350	Sequence 350, App
37	139	7.1	370	14	US-10-173-915-350	Sequence 350, App
38	139	7.1	370	14	US-10-176-706-350	Sequence 350, App
39	139	7.1	370	14	US-10-175-738-350	Sequence 350, App
40	139	7.1	370	14	US-10-175-752-350	Sequence 350, App
41	139	7.1	370	14	US-10-176-482-350	Sequence 350, App
42	139	7.1	370	14	US-10-176-757-350	Sequence 350, App
43	139	7.1	370	14	US-10-176-913-350	Sequence 350, App
44	139	7.1	370	14	US-10-180-552-350	Sequence 350, App
45	139	7.1	370	14	US-10-180-557-350	Sequence 350, App

ALIGNMENTS

RESULT 1
US-10-621-113-4
; Sequence 4, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-113-4

Query Match	100.0%	Score	1950;	DB	15;	Length	371;
Best Local Similarity	100.0%	Pred. No.	8.1e-148;				
Matches	371;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAGPRYPVSQGAALVQIKRLQTFAPSVWSDGSDTFVRSWDEFFQLKKTLETPEVEA	60				
Db	1	MAGPRYPVSQGAALVQIKRLQTFAPSVWSDGSDTFVRSWDEFFQLKKTLETPEVEA	60				
QY	61	GLLRRSDRVLPLKLDAPLGRVGRSGRLARLQLLETYSRRLLATAERVARSPITIGFFA	120				
Db	61	GLLRRSDRVLPLKLDAPLGRVGRSGRLARLQLLETYSRRLLATAERVARSPITIGFFA	120				
QY	121	POPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQLRCLQPFCTQTRDRPFQ	180				

Db 121 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPQ 180
QY 181 AQAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSGPQFCAS 240
Db 181 AQAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSGPQFCAS 240
QY 241 RAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 241 RAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
QY 301 RGGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQGRPGC 360
Db 301 RGGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQGRPGC 360
QY 361 VDSVPHPHTEQ 371
Db 361 VDSVPHPHTEQ 371

RESULT 2

US-10-621-113-6
; Sequence 6, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-113-6

Query Match 99.4%; Score 1937.5; DB 15; Length 376;
Best Local Similarity 98.7%; Pred. No. 8.3e-147; Mismatches 0; Indels 5; Gaps 1;
Matches 371; Conservative 0;

QY 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
Db 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPTI 115
Db 61 GLLRRSDRVLPKLQGAASLDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPTI 120
QY 116 TGFFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTR 175
Db 121 TGFFAPQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTR 180
QY 176 DRPFOAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGP 235
Db 181 DRPFOAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSSGP 240
QY 236 QFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 295
Db 241 QFCASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALL 300
QY 296 SCTGRRGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQ 355
Db 301 SCTGRRGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQ 360
QY 356 RPRGCVDSVPHPHTEQ 371
Db 356 RPRGCVDSVPHPHTEQ 371

Db 361 RPRGCVDSVPHPHTEQ 376
RESULT 3
US-10-621-113-2
; Sequence 2, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-113-2

Query Match 99.2%; Score 1934.5; DB 15; Length 370;
Best Local Similarity 99.7%; Pred. No. 1.4e-146; Mismatches 0; Indels 1; Gaps 1;
Matches 370; Conservative 0;

QY 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 60
Db 1 MAGPRYPVSQGAALVQIKRLQTFAFSVRWSDGSDTFVRRSWDEFRQLKTKLTFPVEA 59
QY 61 GLLRRSDRVLPKLSDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPITGFFA 120
Db 60 GLLRRSDRVLPKLSDAPLLGRVGTSGRLARLQLLETYSRRLATAERVARSPITGFFA 119
QY 121 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPQ 180
Db 120 PQLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPQ 179
QY 181 AQAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSGPQFCAS 240
Db 180 AQAQESLDVLRHPSGWLVENEDRQTAWFPAPYLEEAPGQREGGPSLGSGPQFCAS 239
QY 241 RAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 240 RAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 299
QY 301 RGGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQGRPGC 360
Db 300 RGGDDPAGEARGFPEPSQATAPPTVTPRSPGAIQSRCCCTVTRALERRPRRQGRPGC 359
QY 361 VDSVPHPHTEQ 371
Db 360 VDSVPHPHTEQ 370

RESULT 4

US-10-621-113-8
; Sequence 8, Application US/10621113
; Publication No. US20040091466A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; TITLE OF INVENTION: Regulatory Protein For Nox Enzymes
; FILE REFERENCE: 05501-0202 (43150-287577)
; CURRENT APPLICATION NUMBER: US/10/621,113
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/405,647
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: US 60/396,170

; PRIOR FILING DATE: 2002-07-16										
; NUMBER OF SEQ ID NOS: 11										
; SOFTWARE: PatentIn version 3.1										
; SEQ ID NO 8										
; LENGTH: 375										
; TYPE: PRT										
; ORGANISM: Homo sapiens										
US-10-621-113-8										
Query Match 98.6%; Score 1922; DB 15; Length 375;										
Best Local Similarity 98.4%; Pred. No. 1.4e-145;										
Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;										
Qy	1	MAGPRYPVS	VQGAAL	VOIKRLQ	TFAPF	SVRWS	DGSDT	FVRSW	DFRQLKTKL	KETFPVEA 60
Db	1	MAGPRYPVS	VQGAAL	VOIKRLQ	TFAPF	SVRWS	DGSDT	FVRSW	DFRQLKTKL	KETFPVEA 59
Qy	61	GLLRSDRVLP	PKL-----	LDA	PLGRV	RTSR	GLAR	LQLLETYS	RRLLATAERVARSP	115
Db	60	GLLRSDRVLP	PKLQAS	LDA	PLGRV	RTSR	GLAR	LQLLETYS	RRLLATAERVARSP	119
Qy	116	TGFFAPQ	LDLE	PALPG	SVILPT	PEEQ	PLSRA	AGRLSI	HSLEAQ	SLRCLQPFCTQDTR 175
Db	120	TGFFAPQ	LDLE	PALPG	SVILPT	PEEQ	PLSRA	AGRLSI	HSLEAQ	SLRCLQPFCTQDTR 179
Qy	176	DRPFQA	QAES	LDVLLRHPS	GWMLV	ENEDR	QTAF	FPAPYLEE	AAPGQREGPSLSSGP	235
Db	180	DRPFQA	QAES	LDVLLRHPS	GWMLV	ENEDR	QTAF	FPAPYLEE	AAPGQREGPSLSSGP	239
Qy	236	QFCAS	RAYESS	RADEL	SV	PAGAR	VRVLE	TSDR	GWLCRY	GDRAGLLP 295
Db	240	QFCAS	RAYESS	RADEL	SV	PAGAR	VRVLE	TSDR	GWLCRY	GDRAGLLP 299
Qy	296	SGTGF	RGDD	PAGE	ARGFP	PSQAT	APP	PTVP	TRPS	PGAIQSR
Db	300	SGTGF	RGDD	PAGE	ARGFP	PSQAT	APP	PTVP	TRPS	PGAIQSR
Qy	356	RPRG	CVDS	VPHPT	TEQ 371					
Db	360	RPRG	CVDS	VPHPT	TEQ 375					
RESULT 5										
US-10-437-427-7										
; Sequence 7, Application US/10437427										
; Publication No. US2004009901A1										
; GENERAL INFORMATION:										
; APPLICANT: Rikard Holmdahl										
; APPLICANT: Peter Olofsson										
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase										
; FILE REFERENCE: 11145-024001										
; CURRENT APPLICATION NUMBER: US/10/437,427										
; PRIOR FILING DATE: 2003-05-13										
; PRIOR APPLICATION NUMBER: US 60/380,904										
; PRIOR FILING DATE: 2002-05-13										
; PRIOR APPLICATION NUMBER: US 60/429,609										
; PRIOR FILING DATE: 2002-11-27										
; NUMBER OF SEQ ID NOS: 8										
; SOFTWARE: FastSeq for Windows Version 4.0										
; SEQ ID NO 7										
; LENGTH: 390										
; TYPE: PRT										
; ORGANISM: Homo sapiens										
US-10-437-427-7										
Query Match 15.9%; Score 309.5; DB 15; Length 390;										
Best Local Similarity 25.5%; Pred. No. 2.1e-16;										
Matches 96; Conservative 62; Mismatches 151; Indels 67; Gaps 13;										
Qy	24	FAPF	SVRWS	DGSDT	FVRSW	DFRQLKTKL	KETFP	VBAGL	LRSDRVLP	PKLLDAPLGRVG 83
Db	26	YMF	LKQDL	SEKVVY	RRTEI	YEFH	TKL	KEMF	PIE	BAGAINPENRIIPH-LPAPK
; FILE REFERENCE: 11145-024001										
; CURRENT APPLICATION NUMBER: US/10/437,427										
; PRIOR FILING DATE: 2003-05-13										
; PRIOR APPLICATION NUMBER: US 60/380,904										
; PRIOR FILING DATE: 2002-05-13										
; PRIOR APPLICATION NUMBER: US 60/429,609										
; PRIOR FILING DATE: 2002-11-27										
; NUMBER OF SEQ ID NOS: 8										
; SOFTWARE: FastSeq for Windows Version 4.0										
; SEQ ID NO 7										
; LENGTH: 390										
; TYPE: PRT										
; ORGANISM: Homo sapiens										
US-10-437-427-7										
Query Match 15.8%; Score 307.5; DB 14; Length 390;										
Best Local Similarity 25.3%; Pred. No. 3.1e-16;										
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;										
Qy	24	FAPF	SVRWS	DGSDT	FVRSW	DFRQLKTKL	KETFP	VBAGL	LRSDRVLP	PKLLDAPLGRVG 83
Db	26	YMF	LKQDL	SEKVVY	RRTEI	YEFH	TKL	KEMF	PIE	BAGAINPENRIIPH-LPAPK
Qy	84	RTSR	GLAR	LQLLETYS	RRLLATAERVARSP	TTTGF	APQ	LDLE	PALPG	SVILPT 137
Db	85	RAEN	--RQ	TLTEY	CVSTL	MSLPT	KLIS	RCPL	HLDD	FKVPRDDK--LPTDNQTKK 140
Qy	138	LPTPE	EQPLS	RAAGRLSI	HSLEAQ	SLRCLQ	PFCTQ	DR	TRDRP	FQAQAESLDVLLRHPSG 197
Db	141	MPK	DGK	STATDIT	GPIL-----	Q	TVRA	IADY--	EKTSGSEMA	193
Qy	198	WL	VENED	ROTAF	WPAPYLEE	A--	PGQ	REG	PSL	SGSGQFCASRAYESSRADEL 256
Db	194	WFC	OMKAKR	--GW	IPAS	FLP	LDSP	DE	TEPE	PNY--AGEPYVAIKAYTAVEGDEV 250
Qy	257	AR	VVLE	TS	DRG	WMLC	RYG	DRA	GLLP	AVLLRPEG-----LGA----- 293
Db	251	EAVE	VIH	KLDG	WVIR	KDDVT	GY	FP	SMY	LQSGQDVSAQQRQIKRGAPPRSSIRNV 310
Qy	294	-----	LL	SGT	FR	GGD	-----	DP	AGE	ARGFP-----EP 331

```
Db 311 IHQSRKRLSDAYRNSVRFLOQRROARPGQSPGSPLEBERQTRSKPQPAVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 7

```
US-10-418-036-18
; Sequence 18, Application US/10418036
; Publication No. US20030225117A1
; GENERAL INFORMATION:
; APPLICANT: Gronberg, Alvar
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: 13425-110001
; CURRENT APPLICATION NUMBER: US/10/418,036
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: SE 0201152-6
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/410,626
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-036-18
```

```
Query Match 15.8%; Score 307.5; DB 15; Length 390;
Best Local Similarity 25.3%; Pred. No. 3.1e-16;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
```

```
QY 24 FAFSVRWSDGSDTFVRRSWDFRQLKTLKETFPVEAGLLRRSDRVLKLLDAPLLGRVG 83
Db 26 YNFLVKWQDLSEKVVYRRFTIYEFHKLTKENFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
QY 84 RTRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
Db 85 RAAEN--RQGLTEYCVSTLMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPETYL 140
QY 138 LPTPEQPLSRAAGRLSIHSLAQSLRCLQPCFTQDTRDRPFQQAQESLDVLLRHPGSGW 197
Db 141 MPDQKSTATDITGPILL-----QTYRAIADY--EKTSGSEWALSTGDVVEVKESESGW 193
QY 198 WLVENEDRQTAFWFPAPYLEEA-APGQREGGPGSLGSSGPFQFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASFLPLDSPDETDPENY--AGEPYVAIKAYTAVGDEVSLLEG 250
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQKSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 294 -----LLSGTGFRRGD-----DPAGEARGFP-----EFSQATAPPTVTPRPS 331
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQKSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 311 IHQSRKRLSDAYRNSVRFLOQRROARPGQSPGSPLEBERQTRSKPQPAVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 8

```
US-10-437-427-6
; Sequence 6, Application US/10437427
; Publication No. US20040009901A1
; GENERAL INFORMATION:
; APPLICANT: Rikard Holmdahl
; APPLICANT: Peter Olofsson
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase
; TITLE OF INVENTION: Defects
```

```
; FILE REFERENCE: 11145-024001
; CURRENT APPLICATION NUMBER: US/10/437,427
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/380,904
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/429,609
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-437-427-6
```

```
Query Match 15.8%; Score 307.5; DB 15; Length 390;
Best Local Similarity 25.3%; Pred. No. 3.1e-16;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;
QY 24 FAFSVRWSDGSDTFVRRSWDFRQLKTLKETFPVEAGLLRRSDRVLKLLDAPLLGRVG 83
Db 26 YNFLVKWQDLSEKVVYRRFTIYEFHKLTKEMFPIEAGAINPENRIIPH-LPAPKWFQDQ 84
QY 84 RTRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
Db 85 RAAEN--RQGLTEYCVSTLMSLPTKISRCPHLLDFFKVRPDDLK--LPTDNQTKKPETYL 140
QY 138 LPTPEQPLSRAAGRLSIHSLAQSLRCLQPCFTQDTRDRPFQQAQESLDVLLRHPGSGW 197
Db 141 MPDQKSTATDITGPILL-----QTYRAIADY--EKTSGSEWALSTGDVVEVKESESGW 193
QY 198 WLVENEDRQTAFWFPAPYLEEA-APGQREGGPGSLGSSGPFQFCASRAYESSRADELSPVAG 256
Db 194 WFCQMKAKR-GWIPASFLPLDSPDETDPENY--AGEPYVAIKAYTAVGDEVSLLEG 250
QY 257 ARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
Db 251 EAVEVIHKLDDGWVIRKDDVTGYFFPSMYLQKSGQDVSAQORQIKRGAPRRSSIRNAHS 310
QY 294 -----LLSGTGFRRGD-----DPAGEARGFP-----EFSQATAPPTVTPRPS 331
Db 311 IHQSRKRLSDAYRNSVRFLOQRROARPGQSPGSPLEBERQTRSKPQPAVPPRS 370
QY 332 PGAIQSRCTVTTRAL 347
Db 371 ADLILNRCSESTKRKL 386
```

RESULT 9

```
US-10-767-341-4
; Sequence 4, Application US/10767341
; Publication No. US20040132084A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV-II
; CURRENT APPLICATION NUMBER: US/10/767,341
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 10/109,856
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-767-341-4
```

```
Query Match 15.8%; Score 307.5; DB 16; Length 390;
```

[illegible][illegible]

RESULT 12

```
US-10-437-427-2
; Sequence 2, Application US/10437427
; Publication No. US20040009901A1
; GENERAL INFORMATION:
; APPLICANT: Rikard Olofsson
; APPLICANT: Peter Olofsson
; TITLE OF INVENTION: Autoimmune Conditions and NADPH Oxidase
; TITLE OF INVENTION: Defects
; FILE REFERENCE: 11145-024001
; CURRENT APPLICATION NUMBER: US/10/437,427
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/380,904
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/429,609
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; ORGANISM: Rattus norvegicus
US-10-437-427-2

Query Match
Best Local Similarity 15.7%; Score 306; DB 15; Length 389;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKKTLPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGINPENRVIIPH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITGFAFQPLDLEPALPGSRVILPTPEE 143
D6 85 RAAE--SRQGLTEYFNSLMGLPMKISRCPHLLNFFKVRPDDLK--LPNDSQVKKP---E 137
QY 144 QLSRAAGRLSHSLEA-----QSLCLQPCFCTQDTRDRPFQAQESLDVLLRHPSGWL 199
D6 138 TYLTAKDGNVADINGPILOTYRAIDY--EKSKTEMVTATGDVDDVVEKSSGNWF 195
QY 200 VENEDROTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAGAR 258
D6 196 CQMKTKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVVTKAVAAVEEDSVLSGEA 252
QY 259 VRVLETSRGRWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
D6 253 IEVHKLLDGMVVRKGDITGYFSPMYLQKAGEEITQAQRQIRSRGAPPRKSTIRNAQSI 312
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---PSQATA-PPPTVTPRPSGA 334
D6 313 HORSKRKLSQDYRRNSVRFLQRRRPARPGQSPGSDSKDNFTSTPRAKQPAVPPRPSDL 372
QY 335 IQSRCTVTTRAL 347
D6 373 ILHRCSTESTKRKL 385

RESULT 13
US-10-109-856-2
; Sequence 2, Application US/10109856
; Publication No. US20030166185A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV
; CURRENT APPLICATION NUMBER: US/10/109,856
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-109-856-2

Query Match
Best Local Similarity 14.3%; Score 278.5; DB 16; Length 386;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKKTLPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 26 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGINPENRVIIPH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITGFAFQPLDLEPALPGSRV-----I 137
D6 85 RAAE--ROGLTEYCSLTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDQTKKPEYTL 140
QY 138 LPTPEEQPLSRAAGRLSHSLEAQSRLCLQPCFCTQDTRDRPFQAQESLDVLLRHPSGW 197
D6 141 MPKDGKSTATDITGPILL-----QTYRAINY--EKTSGSEWALSTGDVVEVEKSESGW 193
QY 198 WLVENEDROTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAG 256
D6 194 WFCQMKAKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVVTKAVAAVEEDSVLSLEG 250
QY 257 ARVRVLETSRGRWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
D6 251 EAVEVIHKLDDGW-----KDDVTGYFSPMYLQKSGQVDSQAQRQIKRGAPPRRRSSIRNAHS 306
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---EPSQATAPPPTVTPRPS 331
D6 307 IHORSKRKLSQDYRRNSVRFLQRRRQARPGQSPGSPLEEBERQTSKQPAVPPRPS 366
QY 332 PGAIQSRCTVTTRAL 347
D6 367 ADLILNRCSESTKRKL 382

RESULT 14
US-10-767-341-2
; Sequence 2, Application US/10767341
; Publication No. US20040132084A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001198DIV-II
; CURRENT APPLICATION NUMBER: US/10/767,341
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: 09/820,005
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 10/109,856
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-767-341-2

Query Match
Best Local Similarity 14.3%; Score 278.5; DB 16; Length 386;
Matches 93; Conservative 62; Mismatches 150; Indels 71; Gaps 14;

QY 24 FAFSVRWSGSDTFFVRRSWDEFRQLKKTLPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D6 26 YMFVWKQDLSKVVYRFTIYEFHKLKEMFPIEAGINPENRVIIPH-LPAPRWYDQ 84
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITGFAFQPLDLEPALPGSRV-----I 137
D6 85 RAAE--ROGLTEYCSLTMSLPTKISRCPHLLDFFKVRPDDLK--LPTDQTKKPEYTL 140
QY 138 LPTPEEQPLSRAAGRLSHSLEAQSRLCLQPCFCTQDTRDRPFQAQESLDVLLRHPSGW 197
D6 141 MPKDGKSTATDITGPILL-----QTYRAINY--EKTSGSEWALSTGDVVEVEKSESGW 193
QY 198 WLVENEDROTAWFPAPYLEEA-AFGQREGGPGSLGSGQPCFASRAYESSRADELSVPAG 256
D6 194 WFCQMKAKR-GWIPASYLEPLDSDPEAEDDPNY--AGEPVVTKAVAAVEEDSVLSLEG 250
QY 257 ARVRVLETSRGRWLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
D6 251 EAVEVIHKLDDGW-----KDDVTGYFSPMYLQKSGQVDSQAQRQIKRGAPPRRRSSIRNAHS 306
QY 294 -----LLSGTGRGGD-----DPAGEARGFPE---EPSQATAPPPTVTPRPS 331
D6 307 IHORSKRKLSQDYRRNSVRFLQRRRQARPGQSPGSPLEEBERQTSKQPAVPPRPS 366
QY 332 PGAIQSRCTVTTRAL 347
D6 367 ADLILNRCSESTKRKL 382
```


This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:26:27 ; Search time 40 Seconds
(without alignments)
892.410 Million cell updates/sec

Title: US-10-621-113-4
Perfect score: 1950
Sequence: 1 MAGPRYPVSQGAALVQIKR.....RRQGRPRGCVDSVPHTTEQ 371
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307.5	15.8	390	1 A39249	neutrophil cytosol
2	300.5	15.4	388	2 I54525	leukemia-related p
3	161	8.3	940	2 T00056	hypothetical prote
4	151	7.7	1270	2 T09194	adaptor protein in
5	139.5	7.2	560	1 WFRUM	mullerian inhibiti
6	136.5	7.0	1168	1 MWAXIC	myosin heavy chain
7	128.5	6.6	377	2 S08636	nck protein - huma
8	128	6.6	1249	2 A56511	myosin I myoA - Em
9	125	6.4	1094	2 T13053	dynamlin associated
10	123	6.3	1011	2 T13055	dynamlin associated
11	122.5	6.3	290	2 T42526	hypothetical prote
12	121	6.2	463	2 T34841	probable bifunctio
13	120.5	6.2	1097	2 T31504	hypothetical prote
14	118.5	6.1	2605	2 T18552	saframycin Mx1 syn
15	118	6.1	847	1 A53800	mixed-lineage prot
16	117	6.0	443	2 T27877	hypothetical prote
17	116.5	6.0	780	2 T00366	hypothetical prote
18	116	5.9	1366	2 T35985	probable large Pro
19	115	5.9	470	2 AD0888	Sufi protein (impo
20	115	5.9	512	2 I49552	protein-tyrosine k
21	115	5.9	512	2 E59437	F02569_2 protein [
22	114.5	5.9	303	2 S41754	CRKL protein - hum
23	114.5	5.9	486	2 I49760	LckBP1 protein - m
24	114	5.8	364	2 T35353	hypothetical prote
25	113.5	5.8	639	2 D83591	probable ATP-depen
26	112	5.7	305	1 A49011	c-Crk - chicken
27	111.5	5.7	303	2 S58352	SH2/SH3 adaptor pr
28	111	5.7	304	2 I58394	c-Crk - mouse
29	110.5	5.7	1142	2 T00022	B120 protein - hum

RESULT 1

A39249
neutrophil cytosol factor 1 - human
N;Alternate names: 47K autosomal chronic granulomatous disease protein; multicomponent m
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text change 09-Jul-2004
C;Accession: A35926; A39249; A54067; I59190; A32762; A41385
R;Rodaway, A.R.F.; Teahan, C.G.; Casimir, C.M.; Segal, A.W.; Bentley, D.L.
Mol. Cell. Biol. 10, 5388-5396, 1990
A;Title: Characterization of the 47-kilodalton autosomal chronic granulomatous disease p
A;Reference number: A35926; MUID:90377229; PMID:2398896
A;Accession: A35926
A;Molecule type: mRNA
A;Residues: 1-390 <RD>
A;Cross-references: UNIPROT:PI4598; GB:M55067; GB:M38755; NID:G189050; PIDN:AAA59901.1; I
R;Volpp, B.D.; Nauseef, W.M.; Donelson, J.E.; Moser, D.R.; Clark, R.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9563, 1989
A;Reference number: A39249
A;Contents: erratum
A;Accession: A39249
A;Molecule type: mRNA
A;Residues: 1-390 <VOL>
A;Cross-references: GB:M26193; NID:G189107; PIDN:AAA57209.1; PID:G189108
R;Volpp, B.D.; Nauseef, W.M.; Donelson, J.E.; Moser, D.R.; Clark, R.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 7195-7199, 1989
A;Title: Cloning of the cDNA and functional expression of the 47-kilodalton cytosolic com
A;Reference number: A41385; MUID:89386707; PMID:2550933
A;Contents: annotation
A;Note: the sequence reported has been extensively revised in reference A39249
R;Lomax, K.J.; Leto, T.L.; Nunoi, H.; Gallin, J.I.; Malech, H.L.
Science 245, 409-412, 1989
A;Title: Recombinant 47-kilodalton cytosol factor restores NADPH oxidase in chronic gran
A;Reference number: A32762; MUID:89332501; PMID:2547247
A;Contents: annotation
A;Note: the sequence reported has been extensively revised and now agrees with that show
R;Finan, P.; Shimizu, Y.; Gout, I.; Hsuan, J.; Truong, O.; Butcher, C.; Bennett, P.; Watc
J. Biol. Chem. 269, 13752-13755, 1994
A;Title: An SH3 domain and proline-rich sequence mediate an interaction between two comp
A;Reference number: A54067; MUID:94245680; PMID:8188650
A;Accession: A54067
A;Molecule type: protein
A;Residues: 8-16;44-52;71-77 <PIN>
A;Experimental source: differentiated HL-60 cells
R;Casimir, C.M.; Bu-Gharios, H.N.; Rodaway, A.R.F.;
Proc. Natl. Acad. Sci. U.S.A. 88, 2753-2757, 1991
A;Title: Autosomal recessive chronic granulomatous disease caused by deletion at a dinucl
A;Reference number: I59190; MUID:91187870; PMID:2011585
A;Accession: I59190
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 14-24 <CAS>
A;Cross-references: GB:M60941; NID:G189948; PIDN:AAA60086.1; PID:G189949

hypothetical prote
N-methyl-D-asparta
hypothetical prote
probable membrane
serine/threonine p
hypothetical prote
probable serine/th
DNA helicase RecG
hypothetical prote
rho-GTPase-activat
hypothetical prote
mullerian inhibiti
inositol 1,4,5-tri
nascent polypeptid
translation initia
BOLFI protein - hu

ALIGNMENTS

A>Note: a GT dinucleotide deletion at a GT-GT repeat causes a frameshift after residue 2
C:Comment: This protein is required for activation of the latent NADPH oxidase, which is
ous disease.

C:Genetics:
A:Gene: GDB:NCF1
A:Cross-references: GDB:120222; OMIM:233700
A:Map position: 7q11.23-7q11.23
A:Introns: 24/3
A>Note: the list of introns is incomplete
C:Superfamily: neutrophil cytosol factor 1; SH3 homology
C:Keywords: cytosol; neutrophil
F:163-210/Domain: SH3 homology <SH31>
F:233-280/Domain: SH3 homology <SH32>

Query Match 15.8%; Score 307.5; DB 1; Length 390;
Best Local Similarity 25.3%; Pred. No. 2.1e-14;
Matches 95; Conservative 63; Mismatches 151; Indels 67; Gaps 13;

```
QY 24 FAFSVRWSGSDTFVRRSWDFRQLKTKETFPVVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 26 YMFVKWQDLSEKVVYRRFTEIYEFHKLKEMFPPIEAGAINPENRIIPH-LPAPKWFQDQ 84
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRV-----I 137
DB 85 RAAEN--RQGLTEYCSLTMSLPTKISRCPHLLDFKVRPDDLK--LPTDNQTKKPTVYL 140
QY 138 LPTPEQPLSRAAGRLSHLSLEAQLRCLQPCQTQDTRDRPPQAOAESLDVLLRHPGSGW 197
DB 141 MPKDGKSTATDITGPILL-----QTYRAIADY--BKTSGSEMALSTGDVVEVEKSESGW 193
QY 198 WLVENEDRQTAWFPAPYLEEA-APGQREGGSLGSSGPOFCASRAYESSRADELSVPAG 256
DB 194 WFCQMKAR-GWIPASFLPLDSPDETEDEPNY--AGEPYVAIKAYTAVEGDEVSLLEG 250
QY 257 ARVRYLETSDRGWMLCRYGDRAGLLPAVLLRPEG-----LGA----- 293
DB 251 EAVEVHKLLDGMWVRKDDVTGYFSPMYLQSGQDVSAQOQIKRGAPPRSSIRNAHS 310
QY 294 -----LLSGTGFRRGGD-----DPAGEARGFP-----EFSQATAPPTVTPRPS 331
DB 311 IHQSRKRLSQDQVYRRNSVRFLOQRRRQARPGQSPGSPLEERQTRSKPQPAVPPRPS 370
QY 332 PGAQSRCTVTTRAL 347
DB 371 ADLIILNRCSSETKRKL 386
```

RESULT 2

leukemia-related protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I54525
R:Jackson, S.H.; Malech, H.L.; Kozak, C.A.; Lomax, K.J.; Gallin, J.I.; Holland, S.M.
Immunogenetics 39, 272-275, 1994
A:Title: Cloning and functional expression of the mouse homologue of p47phox.
A:Reference number: I54525; MUID:94164697; PMID:8119734
A:Status: preliminary
A:Accession: I54525
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-388 <RES>
A:Cross-references: GB:L11455; NID:g309422; PIDN:AAA50469.1; PID:g557868
C:Superfamily: neutrophil cytosol factor 1; SH3 homology
F:161-208/Domain: SH3 homology <SH31>

Query Match 15.4%; Score 300.5; DB 2; Length 388;
Best Local Similarity 25.7%; Pred. No. 6.5e-14;
Matches 96; Conservative 55; Mismatches 160; Indels 63; Gaps 11;

```
QY 24 FAFSVRWSGSDTFVRRSWDFRQLKTKETFPVVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DB 24 YMFVKWQDLSEKVVYRRFTEIYEFHKLKEMFPPIEAGIHTENRVIIPH-LPAPRWFDQ 82
```

```
QY 84 RTSRGLARLQLLETYSRRLATAERVARSPITITGFFAPQPLDLEPALPPGSRVILPTBEE 143
DB 83 RAAE--SRQGLTEYFENGLMGLPVKISRCPHLLDFKVRPDDLK--LPTDSQAKP---E 135
QY 144 QPLSRAAGRLSHLSLEA-----QSLRCLQPFCTQDTRDRPPQAOAESLDVLLRHPGSGWWL 199
DB 136 TVLVFKDGNVADITGPILLQTHRAIADY--EKSSGTEMTVATGDVVVDVVEKSESGWVF 193
QY 200 VENEDRQTAWFPAPYLEEA-APGQREGGSLGSSGPOFCASRAYESSRADELSVPAGAR 258
DB 194 CQMKTKR-GWIPASYLEPLDSPDEAEDDPNY--AGEPYVTIKAYAAVEEEMSLSGEA 250
QY 259 VRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGRGDDPAGEA----- 310
DB 251 IEVHKLLDGMWVRKGDITGYFFSWMYLOKAGEEITQQRQIRGRCAPPRSTIRNAOSI 310
QY 311 --RGFPPEPSQAT----- 333
DB 311 HQSRKRLSQDQTYRRNSVRFLOQRRRPGRPRAASTDGTGKNDPSTPRVKPQPAVPPRPSD 370
QY 334 AIQSRCTVTTRAL 347
DB 371 LIHRCSTESTKRKL 384
```

RESULT 3

hypothetical protein KIAA0418 - human
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C:Accession: T00056
R:Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.;
submitted to the EMBL Data Library, October 1997
A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The
A:Reference number: Z14080
A:Accession: T00056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-940 <ISH>
A:Cross-references: UNIPROT:O43302; EMBL:AB007878; NID:d1179754; PIDN:BAA24848.1; PID:d11
A:Experimental source: brain
C:Genetics:
A>Note: KIAA0418

Query Match 8.3%; Score 161; DB 2; Length 940;
Best Local Similarity 24.4%; Pred. No. 0.0012;
Matches 53; Conservative 23; Mismatches 73; Indels 68; Gaps 8;

```
QY 182 QAOESLDVLLRHPSGWMLVENEDRQTAWFPAPYLEEAAPGQREGGSLGSSGPO--FCA 239
DB 23 QAGEVVDVIEKNESGMWPFVSTSEQ-GWVPATYIL-EAQNGTRDDSDINTSKTGESEKVT 80
QY 240 SRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRP----- 288
DB 81 VQPYTSQSKDIBGIEKGVTVEVIRKLEGGWVYRVLGEGWAPASLYLKAKXKDDLPTRKKN 140
QY 289 -----EGLGALLSGTGF-----RGGD---DPA-----GEARG 312
DB 141 LAGPVEITIGNIMEISNLINKKASGDKETPPAEGEGHEAPIAKKEISLPILCNASNGSAVG 200
QY 313 FPE-----PSQATAPPTVTPRPS 332
DB 201 VPDRTVSRLAQSPAVARIAIQAQISSPNLRTTRPPP 237
```

RESULT 4

T09194
adaptor protein interseectin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09194
R:Yamabhai, M.; Hoffman, N.G.; Hardison, N.L.; McPherson, P.S.; Castagnoli, L.; Cesareni,
J. Biol. Chem. 273, 31401-31407, 1998

A;Title: Intersectin, a novel adaptor protein with two eps15 homology and five src homol
A;Reference number: Z16605; MUID:99030416; PMID:9813051
A;Accession: T09194
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1270 <YAM>
A;Cross-references: UNIPROT:O42287; EMBL:AF032118; NID:G2642624; PIDN:AAC73068.1; PID:G2
A;Experimental source: cell type oocyte
C;Function:
A;Description: involved in endocytosis
C;Keywords: endocytosis

Query Match 7.7%; Score 151; DB 2; Length 1270;
Best Local Similarity 27.9%; Pred. No. 0.0085;
Matches 61; Conservative 18; Mismatches 80; Indels 60; Gaps 12;

QY 179 FOQAQESLD-----VLLR--HPSGWLVE-----NEDQTAMFPAPYLEEARPGQGRE- 225
Db 1069 YAATAPEQLTAPGQLILIRKKNPGGWEGELQARGKKRQIGWFFPANYVXLLSPGTNKST 1128

QY 226 -GGPSLGGSGPQFC---ASRAYESSRADELSPAGARVRVLETSDRGWMLCRYGDRAGLL 281
Db 1129 PTEPPKPTSLPTTCQVIGWYDYIAQNDDELAFGKGVINVLNKEDPDWKGELNGHGLP 1188

QY 282 PA---VLLRPGLGALLSGTFRGDDPAGARGFPSPQATAPPTV--PTRSPG--- 333
Db 1189 PSNVKLTLD---MDPSQQFRLGVKPAQ---GTP-----ATGDRPFLFPFRDGPSSLPN 1237

QY 334 -----AIQSRCTVTTRALERPRRQGRPRGCV 362
Db 1238 AFOAPPLSVVMIFRCFTA-----PRFCPD 1262

RESULT 5
WPMUM
mullerian inhibiting factor precursor - human
N;Alternate names: anti-Mullerian hormone; mullerian inhibiting substance (MIS)
C;Species: Homo sapiens (man)
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 09-Jul-2004
R;Accession: A01397
R;Cate, R.L.; Maccalano, R.J.; Hession, C.; Tizard, R.; Farber, N.M.; Cheung, A.; Ninf
an, K.L.; Ragin, R.C.; Manganaro, T.F.; MacLaughlin, D.T.; Donahoe, P.K.
A;Title: Isolation of the bovine and human genes for Mullerian inhibiting substance and
A;Reference number: A90879; MUID:86218082; PMID:3754790
A;Accession: A01397
A;Molecule type: DNA
A;Residues: 1-560 <CAT>
A;Cross-references: UNIPROT:P03971; GB:K03474; NID:G188560; PIDN:AAA98805.1; PID:G386953
C;Comment: Although it does not compete with EGF for receptor binding sites, MIS can inh
C;Comment: For anti-Mullerian hormone type II receptor, see PIR:JCA435.
C;Genetics:
A;Gene: GDB:AMH
A;Cross-references: GDB:118996; OMIM:261550; OMIM:600957
A;Map position: 19p13.3-19p13.3
A;Introns: 138/1; 185/3; 222/1; 275/2
C;Superfamily: inhibin
C;Keywords: cytotoxin; glycoprotein; gonadal differentiation; hormone; testis
F;12/Domain: signal sequence #status predicted <SIG>
F;22-25/Domain: propeptide #status predicted <PRO>
F;26-560/Product: mullerian inhibiting factor #status predicted <MAT>
F;64,329/Binding site: carboxylate (Asn) (covalent) #status predicted
F;462-526,488-557,492-559/disulfide bonds: #status predicted
F;525/disulfide bonds: interchain #status predicted

Query Match 7.2%; Score 139.5; DB 1; Length 560;
Best Local Similarity 28.6%; Pred. No. 0.022;
Matches 93; Conservative 20; Mismatches 147; Indels 65; Gaps 18;

QY 89 LARLQLETVYRRLTAAERVARSPRI--TGFFAPQPLDLEPALP--PGSRVIL---PTP 141
Db 6 LTSALVALGALLGTEALREAPAVGTSGLIFREDLWPPGIPQEPFLCLVALGDSNG 65

A;Title: myosin heavy chain IC - Acanthamoeba castellanii
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Acanthamoeba castellanii
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
R;Accession: A33891; C34448; A24146
R;Jung, G.; Korn, E.D.; Hammer III, J.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 6720-6724, 1987
A;Title: The heavy chain of Acanthamoeba myosin IB is a fusion of myosin-like and non-my
A;Reference number: A33891; MUID:88016163; PMID:3477803
A;Accession: A33891
A;Molecule type: DNA
A;Residues: 1-1168 <JUN>
A;Cross-references: UNIPROT:P10569; GB:J02974; NID:G155624; PIDN:AAA27707.1; PID:G155625
A;Note: this gene and protein are called MIB in this paper
R;Brzecka, H.; Lynch, T.J.; Martin, B.; Korn, E.D.
J. Biol. Chem. 264, 19340-19348, 1989
A;Title: The localization and sequence of the phosphorylation sites of Acanthamoeba myosi
A;Reference number: A34448; MUID:90037074; PMID:2530230
A;Accession: C34448
A;Molecule type: protein
A;Residues: 308-314, 'X', 316-329 <BRZ>
C;Comment: In this protein, the coiled-coil rod-like region found in many myosin heavy ct
he protein is globular and does not self-associate into filaments.
C;Genetics:
A;Gene: MTC
A;Introns: 1/3; 37/3; 60/2; 100/2; 153/3; 179/3; 208/2; 242/3; 287/3; 321/3; 371/3; 428/3;
C;Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolo
C;Keywords: actin binding; ATP; hydrolase; nucleotide binding; P-loop; phosphoprotein; t
F;10-653/Domain: myosin motor domain homology <MMOT>
F;101-108/Region: nucleotide-binding motif A (P-loop)
F;543-564/Region: actin binding #status predicted
F;671-1168/Domain: carboxyl-terminal <CTD>
F;675-883/Region: basic
F;923-978/Region: alanine/glycine/proline-rich
F;983-1030/Domain: SH3 homology <SH3>
F;1034-1168/Region: alanine/glycine/proline-rich
F;107/Binding site: ATP (lys) #status predicted
F;311/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 7.0%; Score 136.5; DB 1; Length 1168;
Best Local Similarity 24.4%; Pred. No. 0.081;
Matches 71; Conservative 18; Mismatches 85; Indels 117; Gaps 14;

QY 124 LDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSLRCLQPCTQDTRDRPTQAOA 183
Db 884 VSVAPGLPSSAPNTQAPQE-----TSGGAS-----FTVAE 914

QY 184 QESLDVLL-----RHPSGWLVENEDQTAMFPAPYLEEARPGQGRE 228
Db 915 QSYKDQILKAGGGGGGRGGRGSPSG-----AVSRP-----SPG-GGGGGP 957

QY 229 S-----LGSSGQFCAS-----RA---YESSRADELSVPAGARVRVLETSDRGWMLCR 273
Db 958 SPFGGRPSFGPAAASAPGPEQARALYDFAENDELTFNEGAVVTVINKSNPDWGE 1017
QY 274 YGDRAGLLPAVLL-----RP-----EGLGALLSGTG---FRGGDDPA 307
Db 1018 LNGQGVFPASIVELIPRAAAPAGPSGGPRPAPGGKSGRAAPMGPGPMRGCGPAPG 1077
QY 308 GEAR-GFPEPSQATAPPVTRPSPGAIQSRCCCTVTRRALERRPRQR 357
Db 1078 GPGRGAGPPGAGRAGPGGRCMPAGG-----AAPRGAP 1114
RESULT 7
S08636
nck protein - human
N:Alternate names: src-related protein
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S08636
R:Lehmann, J.M.; Riethmuller, G.; Johnson, J.P.
Nucleic Acids Res. 18, 1048, 1990
A:Title: Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homol
A:Reference number: S08636; MUID:90192089; PMID:2107526
A:Accession: S08636
A:Molecule type: mRNA
A:Residues: 1-377 <LEH>
A:Cross-references: UNIPROT:P16333; EMBL:X17576; NID:g35014; PIDN:CAA35599.1; PID:g35015
F:9-56/Domain: SH3 homology <SH31>
F:113-160/Domain: SH3 homology <SH32>
F:197-247/Domain: SH3 homology <SH33>
F:282-371/Domain: SH2 homology <SH2>
Query Match 6.6%; Score 128.5; DB 2; Length 377;
Best Local Similarity 24.6%; Pred. No. 0.083;
Matches 44; Conservative 20; Mismatches 52; Indels 63; Gaps 6;
QY 179 FOAQAESLDV-----LLRHPSQWLVENEDROTANFPAPYLE-----E 217
Db 13 YVAQOEQLDIKKRLMLLDDSKSWVRVNSMKTGFVPSNYVERKNSARKASIVKNLK 72
QY 218 AAPGGR-EGGPSLGGSPQFCASRA-----YESSRADELSV 253
Db 73 DTLGIGKVRKSPVDS-----ASPADDSFVDPGBRLYDLNMPATVKNYMAERDELSL 127
QY 254 PAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRPEGLGALLSGTGFRGGDDPAGEARG 312
Db 128 IKGTVIVWEKSCDWGRGSYNGQVGFPSNVYTEE-----GDSPLGDHVG 173
RESULT 8
A56511
myosin I myoA - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C>Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 09-Jul-2004
C:Accession: A56511
J. Cell Biol. 128, 577-587, 1995
A:Title: myoA of Aspergillus nidulans encodes an essential myosin I required for secreti
A:Reference number: A56511; MUID:95164560; PMID:7860631
A:Accession: A56511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1249 <MCG>
A:Cross-references: UNIPROT:Q00647; GB:U12427; NID:G525321; PIDN:AAA67877.1; PID:G525322
C:Genetics:
A:Gene: myoA
C:Superfamily: protozoan myosin heavy chain IB; myosin motor domain homology; SH3 homolo
C:Keywords: nucleotide binding; P-loop
F:53-716/Domain: myosin motor domain homology <MMOT>
F:143-150/Region: nucleotide-binding motif A (P-loop)
F:1081-1130/Domain: SH3 homology <SH3>

Query Match 6.6%; Score 128; DB 2; Length 1249;
Best Local Similarity 23.4%; Pred. No. 0.35;
Matches 62; Conservative 25; Mismatches 92; Indels 86; Gaps 12;
QY 128 PALPFGSRVILPTPE-----EQPLSRAAGRLSIHSLEA-----QSURC 165
Db 970 PGEFPNS-VSKPTPRGQVAARPVTK--GKLLRLAVQAVARNPWLPLDYQSGLYHSPRL 1026
QY 166 LQPCTQDTRDRPFOAQAESLDVLLRHPSQWLVENEDROTANFPAPYLEEAPQGORE 225
Db 1027 KQPRNRHQRDPDFLNQWP-----LQHPI-----HVLHLLPPOGHH 1063
QY 226 GGPSL-----GSSGPOFC-ASRAYESSRADELSVPAGARVRVLETSDRGWMLCRYGDRA- 278
Db 1064 --PRLLPRPPAAAGPKAKALYDFSSDNNNGMLSIAGQIVKSGEGNWLCNNLETS 1121
QY 279 -GLLPAVLLRPEGLGALLSGTGFRGGDDPAGEARGFPPEPSQATAPPVTRPSPGAIQS 337
Db 1122 QGWTPEAVL-----EEQVAPTFKPAPPPPPVAPRASPAPVNG 1159
QY 338 RCTVTRRAL-----ERRPRQR 356
Db 1160 SAAVAATAAKAAAPPPPAKRPNMAGR 1184
RESULT 9
TI3053
dynamain associated protein isoform Dapl60-1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: TI3053
R:Roos, J.; Kelly, R.B.
J. Biol. Chem. 273, 19108-19119, 1998
A:Title: Dapl60, a neural-specific Eps15 homology and multiple SH3 domain-containing prot
A:Reference number: Z17594; MUID:98334647; PMID:9668096
A:Accession: TI3053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1094 <ROO>
A:Cross-references: UNIPROT:O61618; EMBL:AF0533957; NID:g2984714; PID:g2984715; PIDN:AAAC3
C:Genetics:
A:Gene: Dapl60
A:Cross-references: FlyBase:FBgn0023388
Query Match 6.4%; Score 125; DB 2; Length 1094;
Best Local Similarity 23.9%; Pred. No. 0.49;
Matches 55; Conservative 33; Mismatches 86; Indels 56; Gaps 10;
QY 106 AERVARSPTTGFPAQPOPLDEPA-LPPGSRVILPTPE---QPLSRAA---GRLSIHS 157
Db 881 SEQPISSPGVGAERAEHDLDEVSQINTQSTQSSSEPAESYSRMPSRTSSMTFGNRAKRS 940
QY 158 LEAQSRLCLQPFCTQDTRDRPFOAQAESLD-----VLLRH---PSGWM-----LVENEDR 205
Db 941 EIAQVI-----APVEATSTEQLSLTRGQLIMIRKKTDSGWEGELQAKGRRR 987
QY 206 QTAWFPAPYLEEAPAGQREGPSLGGSP-----QFCASRAYESSRADEL 251
Db 988 QIGWFPATYVKVL-----QGRNSGRNTPVSGSRIEMTEQILDKVIALYPYKAQNDDEL 1041
QY 252 SVPAGARVRVLETSDRGWMLCRYGDRAGLLPAVLLRP---EGLGALLSGT 298
Db 1042 SFDKDDIISVLGRDEPEWWRGELNGLSLGFLFNSVYGVPTSGKPAKANGT 1091
RESULT 10
TI3055
dynamain associated protein isoform 160-2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: TI3055
R:Roos, J.; Kelly, R.B.

RESULT 12

T34841
Probable bifunctional synthase /transferase - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C/Accession: T34841
R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrall, B.G.; Rajan
submitted to the EMBL Data Library, February 1999
A/Reference number: Z21559
A/Accession: T34841

A;Reference number: Z21559
A;Accession: T34841

A; Status: preliminary; translated from GB/EMBL/DBJ

A; Cross-references: UNIPROT: Q9Z5B5; EMBL: AL035478; A. Experimental source: strain A3(2)

C;Superfamily: hypothetical protein b3052

Matches 82; Conservative 34; Mismatches 13

Db 67 EVTLIAGVGEDPAGLALRELLAPWLKLI-----

Db 107 RVLAQDRPVVRLDRGGGRVREATDEARDALGCARAVL

D**b** 162 RPPLVWDPHPRGGPPVPGTRLVTPAEKEAHGF-----

Db 211 AAALVRDWRVAAVTTLGSRGALLSYGEHPLLVPAPAA

Db 271 ---GALV-GEAVEGA---VGAATAFVAAGGAAAAPPAA

DB 324 AEHGTVAAGGCFD¹LLHAGHVGLLQAARRLGDC¹LVVC

Db 384 VRVLRALACVDAV 396

T31504
hypothetical protein Y116A8C.36 - Caenorhabditis elegans

C; Accession: T31504
R: McMurray, A.

A;Accession: T31504
A;Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-1097 <W1>
A:Cross-references: UNIPROT:Q9U2T9; EMBL:AL117204; PIDN:CA855138.1; CESP:Y116A8C.36
A:Experimental source: clone Y116A8C
C:Genetics:
A:Gene: CESP:Y116A8C.36
A:Introns: 50/3; 245/3; 411/3; 486/3; 697/1; 820/3; 926/2

Query Match 6.2%; Score 120.5; DB 2; Length 1097;
Best Local Similarity 19.9%; Pred. No. 1;
Matches 57; Conservative 35; Mismatches 107; Indels 87; Gaps 10;

QY 104 ATAERVARSPITITGFAPQDLLEALPPGSRVI-----LPTPEEOPL 146
DB 727 AFVEIAAVPTPGG--DPPIQNPMPNTPSSVDQIGVKAARKAIAAMGUTEGAPPA 784
QY 147 SRAAGRLSHSLEAQSRLCLPQCTQDRDRPFOAQAESLDVLLRHPSGMWLVENEDRQ 206
DB 705 SSAPAAAVIS-----QCIAQFQWRARNEEDLSFPAKGTIEVLEKQEMK-WKGRNPAGE 837
QY 207 TAWFPAPYLEE-----AAPQREGGPGSLGS-----SGPQ-- 236
DB 838 IGWFPKSYKVGATTSTTTPIVSPSKASAGAPGAAGAAQAYDVVPSDVTLOASETAPQOQ 897
QY 237 PCASRAVESSRADELSVPAGARVRLVETSDRGWLCRCVDRAGLLPAVLLRPEGLGALL 295
DB 898 LVTVIYDFEAVETTDLAHLVGTITILVLEKNDE-WWKGRCNGREGIPPA----- 944
QY 296 SGTGFRGGDDPAGEARGPEPSQATAPPTVTR-PSPCAISQSCC 340
DB 945 -----NVEISVQAGDPTPTQAPTAAAPTIVLC 974

RESULT 14
T18552
saframycin Mx1 synthetase A - Myxococcus xanthus
C:Species: Myxococcus xanthus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18552
R:Pospiech, A.; Bietenhader, J.; Schupp, T.
Microbiology 142, 741-746, 1996
A:Title: Two multifunctional peptide synthetases and an O-methyltransferase are involved
A:Reference number: Z18967; MUID:97090395; PMID:8936303
A:Accession: T18552
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2605 <POS>
A:Cross-references: UNIPROT:Q50858; EMBL:U24657; NID:g1171127; PID:g1171129; PIDN:AAC441
C:Genetics:
A:Gene: safA

C:Keywords: carrier protein
F:535-978/Domain: acetate-CoA ligase homology <ACL1>
F:997-1065/Domain: acyl carrier protein homology <ACP1>
F:1643-2091/Domain: acetate-CoA ligase homology <ACL2>
F:2110-2178/Domain: acyl carrier protein homology <ACP2>

Query Match 6.1%; Score 118.5; DB 2; Length 2605;
Best Local Similarity 24.9%; Pred. No. 3.7;
Matches 82; Conservative 33; Mismatches 105; Indels 109; Gaps 15;

QY 59 EAGLLRRSDRLVKLLDAPLLGRVGRTS-----RGL-----ARLQLLETYSRLLATAE 107
DB 611 EAGLVGEGDANLEPVADAAQLAYVLYTSGSSGRPKGVWVSHGALANFLTITWAREPGLRAE 670
QY 108 RVARSPTITGFAPQDLLEALPPGSRVILPTPEQPLSRA-AGRLSHSLEAQSRLCL 166
DB 671 DVLAAVT-TFSPDIAALELYLPLVQGVWVATREQAADGRALSGVLARH----- 719
QY 167 QPFCTQDRDRPFOAQAESLDVLLRHPSGMWLVENEDRQTAWFPAPYLEEAPQCGREG 226
DB 720 -----GVTVMQATPATWRML-----ADA 737
QY 227 GPSLSSGGPQFCASRAYESSRADELSVPAGARV-RVLETSDRGWMLCR-----YGDAGLL 281

DB 738 GGAPCTGFTVLCCGALPQDLADALTA-NGARVWNLVYGTETTTVWVSCRKRLGAGDRVSLG 796
QY 282 PA-----VL---LRPEGLGALLSGTGRGDDPAGEARGF-----PEPSQA 319
DB 797 GALTGNTSVHVLDPDLRPVPVG--LAGELFIGG---SGVARGYWGRRPSITAEFRFVDPFSA 851
QY 320 TAPPPTVTRPSPGAIQSRCCCTVTRALE 348
DB 852 -----RPGARLYRTGDLVRRRVYD 869

RESULT 15
A53800
mixed-lineage protein kinase (EC 2.7.1.1-) 3 - human
N:Alternate names: protein kinase PTKI; protein kinase SPRK
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A53800; I58395
R:Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-containing protein kinase
A:Reference number: A53800; MUID:94253068; PMID:8195146
A:Accession: A53800
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-847 <GAL>
A:Cross-references: UNIPROT:Q16584; GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028
R:Jing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, N.J.
Oncogene 9, 1745-1750, 1994
A:Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain
A:Reference number: I58395; MUID:94239754; PMID:8183572
A:Accession: I58395
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-847 <RES>
A:Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296
C:Genetics:
A:Gene: GDB:MLK3; PTKI; SPRK
A:Cross-references: GDB:I34755; OMIM:600050
A:Map position: 11q13.1-11q13.3
C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein kinase
F:48-100/Domain: SH3 homology <SH32>
F:115-383/Domain: protein kinase homology <KIN>
F:123-131/Region: protein kinase ATP-binding motif
F:403-424/Region: leucine zipper motif
F:438-459/Region: leucine zipper motif
F:468-482/Region: basic

Query Match 6.1%; Score 118; DB 1; Length 847;
Best Local Similarity 34.3%; Pred. No. 1.1;
Matches 35; Conservative 9; Mismatches 38; Indels 20; Gaps 4;

QY 221 GGGP-EGGP-SLGSSGPGQFCASRAYESSRADELSVPAGARVRLVLE-----TSDRGWMLCR 273
DB 28 GGRPEGSFKAAGYANPVWTALTIFYEPCQDELARKGDRVEVLSRDAAISGDEGWWAGQ 87
QY 274 YGDRAGLLPAVLLRPEGLGALLSGTGRGDDPAGEARGFPE 315
DB 88 VGGQVGIFP-----SNVSRGGGPPPPCEVASFOE 116

Search completed: May 28, 2005, 07:41:21
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2005, 07:27:23 ; Search time 169 Seconds
(without alignments)
1124.150 Million cell updates/sec

Title: US-10-621-113-4

Perfect score: 1950

Sequence: 1 MAGPRYPVSVQGAALVOIKR.....RRQGRPRGCVDSVPHTTEQ 371

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1950	100.0	371	2	Q8NFA3	Q8nfa3 homo sapien
2	1937.5	99.4	376	2	Q8NFA2	Q8nfa2 homo sapien
3	1934.5	99.2	370	2	Q96B73	Q96b73 homo sapien
4	1922	98.6	375	2	Q86YM1	Q86ym1 homo sapien
5	1219.5	62.5	349	2	Q8VCN2	Q8vcn2 mus musculus
6	1219.5	62.5	349	2	Q8BH41	Q8bh41 mus musculus
7	927	47.5	253	2	Q76EW0	Q76ew0 cavia porce
8	857.5	44.0	239	2	Q9D747	Q9d747 mus musculus
9	321	16.5	423	2	Q7T042	Q7t042 fugu rubrip
10	307.5	15.8	330	1	NCF1_HUMAN	P14598 homo sapien
11	307	15.7	389	2	Q811Y3	Q811y3 rattus norv
12	306.5	15.7	390	2	Q7TMH0	Q7tmh0 mus musculus
13	306	15.7	389	2	Q99M65	Q99m65 rattus norv
14	303.5	15.6	330	1	NCF1_MOUSE	Q99014 mus musculus
15	302	15.5	391	2	Q95MN0	Q95mn0 oryctolagus
16	300.5	15.4	392	1	NCF1_BOVIN	Q77774 bos taurus
17	299.5	15.4	382	2	Q95L71	Q95l71 bison bison
18	298	15.3	391	2	Q9N0E8	Q9n0e8 tursiops tr
19	278	14.3	310	2	Q9JK56	Q9jk56 rattus norv
20	246.5	12.6	1124	2	Q89032	Q89032 mus musculus
21	245.5	12.6	563	2	Q6DC19	Q6dc19 brachydanio
22	240	12.3	471	2	Q8BIC6	Q8bic6 mus musculus
23	164	8.4	805	2	Q7S016	Q7s016 neurospora
24	163	8.4	174	2	Q76EV8	Q76ev8 cavia porce
25	161	8.3	989	2	Q43302	Q43302 homo sapien
26	157	8.1	1228	2	Q6C7C0	Q6c7c0 yarrowia li
27	151	7.7	1270	1	ITN1_XENLA	Q42287 xenopus lae
28	148	7.6	1031	2	Q9H462	Q9h462 homo sapien
29	144.5	7.4	690	2	Q86WN1	Q86wn1 homo sapien
30	144.5	7.4	701	2	Q9NXX8	Q9nxx8 homo sapien
31	139.5	7.2	560	1	MIS_HUMAN	P03971 homo sapien

RESULT 1
Q8NFA3 ID Q8NFA3 PRELIMINARY; PRT; 371 AA.
AC Q8NFA3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Regulatory protein NOX01-beta (NADPH oxidase organizer 1).
GN Name=NOX01;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14617635; DOI=10.1074/jbc.M305968200;
RA Cheng G., Lambeth J.D.;
RT "NOX01, regulation of lipid binding, localization, and activation of
RT Nox1 by the Phox homology (PX) domain."
RL J. Biol. Chem. 279:4737-4742(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22716264; PubMed=12716910; DOI=10.1074/jbc.M212856200;
RA Takeya R., Ueno N., Kami K., Taura M., Kohjima M., Izaki T., Nunoi H.,
RA Sumimoto H.;
RT "Novel human homologues of p47phox and p67phox participate in
RT activation of superoxide-producing NADPH oxidases."
RL J. Biol. Chem. 278:25234-25246(2003).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=12473664; DOI=10.1074/jbc.C200613200;
RA Banfi B., Clark R.A., Steger K., Krause K.-H.;
RT "Two Novel Proteins Activate Superoxide Generation by the NADPH
RT Oxidase NOX1."
RL J. Biol. Chem. 278:3510-3513(2003).
CC - SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AF532984; AAM97926.1; -
DR EMBL; AB097667; BAC76711.1; -
DR EMBL; AF539796; AAN75141.1; -
DR HSSP; P14598; IKQ6.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3.1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS0002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 371 AA; 40796 MW; 39E13C3FF491797E CRC64;

Query Match 100.0%; Score 1950; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 4.3e-121;

Q6ux75 homo sapien
Q55033 mus musculus
Q8bq28 mus musculus
Q6dbu6 brachydanio
Q24218 drosophila
Q8ipw2 drosophila
Q9vpul drosophila
P10569 acanthamoeb
Q61080 acanthamoeb
P79956 xenopus lae
P79956 xenopus lae
Q43639 homo sapien
Q7ah03 neurospora
Q66115 xenopus tro
Q803a2 brachydanio

```
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
QY 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPPTTIGPFA 120
Db 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSPPTTIGPFA 120
QY 121 PQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPPQ 180
Db 121 PQPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTRDRPPQ 180
QY 181 AQAQESLDVLLRHPGSMWLVENEDRQTAWFPAPYLEEAPGQREGGSLGSSGPFQFAS 240
Db 181 AQAQESLDVLLRHPGSMWLVENEDRQTAWFPAPYLEEAPGQREGGSLGSSGPFQFAS 240
QY 241 RAYESSRADELSPAGARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
Db 241 RAYESSRADELSPAGARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTGF 300
QY 301 RGGDDPAGEARGFPESQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQGRPGC 360
Db 301 RGGDDPAGEARGFPESQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQGRPGC 360
QY 361 VDSVPHPTTEQ 371
Db 361 VDSVPHPTTEQ 371
RESULT 2
Q8NFA2 PRELIMINARY; PRT; 376 AA.
AC Q8NFA2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Regulatory protein NOXO1-gamma.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lambeth J.D., Cheng G.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2002)
CC -!- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AF532985; AAM97927.1; -.
DR HSSP; P14598; 1K06.
DR Genew; HGNC:19404; NOXO1.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3 1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS50195; PX; 1.
DR PROSITE; PS50002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 376 AA; 41252 MW; 2E9525A8BADD360 CRC64;
Query Match 99.4%; Score 1937.5; DB 2; Length 376;
Best Local Similarity 98.7%; Pred. No. 2.9e-120;
Matches 371; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
Db 1 MAGPRYPVSVOGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFRQLKTKLKEFPVEA 60
QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSP 115
Db 61 GLLRRSDRVLPKLQQAASLDAPLLGRVGRVTSRGLARLQLLETYSRRLLATAERVARSP 120
```

```
QY 116 TGFAPOPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
Db 121 TGFAPOPLDLEPALPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 180
QY 176 DRPFOAQAESLDVLLRHPGSMWLVENEDRQTAWFPAPYLEEAPGQREGGSLGSSGP 235
Db 181 DRPFOAQAESLDVLLRHPGSMWLVENEDRQTAWFPAPYLEEAPGQREGGSLGSSGP 240
QY 236 QFCASRAYESSRADELSPAGARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 295
Db 241 QFCASRAYESSRADELSPAGARVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 300
QY 296 SCTGPRGDDPAGEARGFPESQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQ 355
Db 301 SCTGPRGDDPAGEARGFPESQATAPPTVTRPSPGAIQSRCTVTRRALLERPRRQ 360
QY 356 RPRGCVDSVPHPTTEQ 371
Db 361 RPRGCVDSVPHPTTEQ 376
RESULT 3
Q96B73 PRELIMINARY; PRT; 370 AA.
AC Q96B73;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TremBLrel. 28, Last annotation update)
DE NADPH oxidase organizer 1, isoform a (NADPH oxidase regulatory
DE protein) (Regulatory protein NOXO1-alpha).
GN Name=NOXO1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Caavaant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22651106; PubMed=12657628; DOI=10.1074/jbc.M301289200;
RA Geiszt M., Lekstrom K., Witta J., Leto T.L.;
RT "Proteins homologous to p47phox and p67phox support superoxide
RT production by NAD(P)H oxidase 1 in colon epithelial cells."
RJ J. Biol. Chem. 278:20006-20012 (2003).
```



```

[4]
RN SEQUENCE FROM N.A.
RA Lambeth J.D., Cheng G.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; BC015917; AH1517.1; -.
DR EMBL; AX255768; AAPI3479.1; -.
DR EMBL; AF532983; AAM97925.1; -.
DR HSSP; P14598; 1KQ6.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS50002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 370 AA; 40668 MW; 1EB8DCB8BC50551F CRC64;

Query Match 99.2%; Score 1934.5; DB 2; Length 370;
Best Local Similarity 99.7%; Pred. No. 4.5e-120;
Matches 370; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPRLKTKLKEPVEA 60
Db |||||
QY 61 GLLRRSDRVLPKLLDAPLLGRVGRVTSRGLARLQLLETYSRRLLTAAERVARSPPTI 120
Db |||||
QY 121 PQLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPFQ 180
Db |||||
QY 120 PQLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQTRDRPFQ 179
QY 181 AQAQESLDVLLRHPSGWLWENEDROTAWFPAPYLEEAPGQREGGSPISGSGPOFCAS 240
Db |||||
QY 180 AQAQESLDVLLRHPSGWLWENEDROTAWFPAPYLEEAPGQREGGSPISGSGPOFCAS 239
QY 241 RAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTG 300
Db |||||
QY 240 RAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALLSGTG 299
QY 301 RGDDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCCCTVTRALERRRRPQGRPGC 360
Db |||||
QY 300 RGDDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCCCTVTRALERRRRPQGRPGC 359
QY 361 VDSVPHPTTEQ 371
Db |||||
QY 360 VDSVPHPTTEQ 370

RESULT 4
Q86YM1 PRELIMINARY; PRT; 375 AA.
ID Q86YM1
AC Q86YM1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Regulatory protein NOX01-delta.
GN Name=NOX01;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A.
RA Cheng G., Lambeth J.D.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AV191359; AAO38665.1; -.
DR HSSP; P14598; 1KQ6.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.

us-10-621-113-4.rup
SMART; SM00326; SH3; 2.
DR PROSITE; PS50002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 375 AA; 41124 MW; 85E41025A14AE80 CRC64;

Query Match 98.6%; Score 1922; DB 2; Length 375;
Best Local Similarity 98.4%; Pred. No. 3.1e-119;
Matches 370; Conservative 0; Mismatches 0; Indels 6; Gaps 2;

QY 1 MAGPRYPVSVQGAALVQIKRLQTFAPSVRWSDGSDTFVRRSWDEFPRLKTKLKEPVEA 60
Db |||||
QY 61 GLLRRSDRVLPKL-----LDAPLLGRVGRVTSRGLARLQLLETYSRRLLTAAERVARSPPTI 115
Db |||||
QY 116 TGFFAPQPLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 175
Db |||||
QY 120 TGFFAPQPLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLAQSLRCLQPFCTQDTR 179
QY 176 DRPFOAQAESLDVLLRHPSGWLWENEDROTAWFPAPYLEEAPGQREGGSPISGSGP 235
Db |||||
QY 180 DRPFOAQAESLDVLLRHPSGWLWENEDROTAWFPAPYLEEAPGQREGGSPISGSGP 239
QY 236 QFCASRAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 295
Db |||||
QY 240 QFCASRAYESSRADELSPAGARVVRVLETSDRGWLCRYGDRAGLLPAVLLRPEGLGALL 299
QY 296 SGTGFRGGDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCCCTVTRALERRRRPQGR 355
Db |||||
QY 300 SGTGFRGGDDPAGEARGPPEPSQATAPPTVTRPSPGAIQSRCCCTVTRALERRRRPQGR 359
QY 356 RPRGCVDSVPHPPTTEQ 371
Db |||||
QY 360 RPRGCVDSVPHPPTTEQ 375

RESULT 5
Q8VCM2 PRELIMINARY; PRT; 349 AA.
ID Q8VCM2
AC Q8VCM2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Noxol protein.
GN Name=Noxol;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler K.G.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalobos D., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richardson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```

```
RT and mouse cDNA sequences." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; BC019525; AAH19525.1; -.
DR HSP; O89100; 10EB.
DR MGD; MGI:1919143; Noxol.
DR GO; GO:0016176; P:superoxide-generating NADPH oxidase activat. . .; IDA.
DR GO; GO:0006801; P:superoxide metabolism; IDA.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3_1; 2.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS00002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 349 AA; 38841 MW; 5008801247454101 CRC64;
Query Match 62.5%; Score 1219.5; DB 2; Length 349;
Best Local Similarity 67.4%; Pred. No. 8.1e-73;
Matches 242; Conservative 31; Mismatches 69; Indels 17; Gaps 4;
QY 1 MAGPRYPVSQGAALVQIKRQTFAFVSVDGSDTFVRRSWDEPRLQKTLKTFPVEA 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MASPRHPVSAHAVALQMDRLQTFAFVSVDGSDTFVRRSWDEPRLQKTLKTFPVEA 60
QY 61 GLLRRSDVLPKLLDAPILGRVGRSGLARLQILETYSRLRLATAERVARSPITGFA 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 GLLRRSEQVLPKLPAPLLTRGHTGRGLVRLRLDLYVQALLATSEHLRSSALHGFFV 120
QY 121 POPLDLEALPPGSVILPTPEEQPLRSRAAGRLSIHSLEAOSLRCLQPCTQDTRDRPFQ 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 PKPLDLEMLPGLSVILPTPEE-PLSQPRSLDIHSLEAQSPICVQPFHLLDIRDRPF 179
QY 181 AQAQESDLVLRHPSGWWLVNEDRQTAFWPAPYLEEAAPQGRGSPSLGSSGPFQFAS 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 TKAQEIIDLRLHPSGWWLVNEDKQVAFWPAPYLEEVATCQGESGLALQSGRQFCTT 239
QY 241 RAYESSRADELSPAGAVRVLETSDRGWMLCRYGDRAGLAPVLLRPEGLGALLSGTGF 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 QAYEGSRDELSPVGARVHVLETSDRGWMLCRYNGRTGLLPAMSLQPEGLGSLGRPGF 299
QY 301 --RGGDDPAGARGPPEPSQATAPPTVTPTRPSPGAIQSRCTVTTRAL-----ERRPR 352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 PDSAGADKVAEDRTI-----PPVVTPTPCMSAIQSRCCSITRRALAQEQGTTRVPR 349
RESULT 6
Q8BH41 PRELIMINARY; PRT; 349 AA.
AC Q8BH41;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NADPH oxidase organizer 1 (Mus musculus 2 days neonate thymus thymic
DE cells cDNA. RIKEN full-length enriched library, clone.E430007K11
DE product:231034C04RIK PROTEIN (SNX28) homolog).
GN Name=Noxol;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX PubMed=12473664; DOI=10.1074/jbc.C200613200;
RA Banfi B., Clark R.A., Steger K., Krause K.-H.;
RT "Two Novel Proteins Activate Superoxide Generation by the NADPH
```

```
RT Oxidase NOX1." ;
RL J. Biol. Chem. 278:3510-3513 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning." ;
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs." ;
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes." ;
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Iizawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer." ;
RL Genome Res. 10:1757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numataki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito D., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki R., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; AF539797; AA075142.1; -.
DR EMBL; AK088226; BAC40222.1; -.
DR HSP; O89100; 10EB.
DR MGD; MGI:1919143; Noxol.
DR GO; GO:0016176; P:superoxide-generating NADPH oxidase activat. . .; IDA.
DR GO; GO:0006801; P:superoxide metabolism; IDA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3_1; 2.
```

```
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00312; PX; 1.
DR SMART: SM00326; SH3; 2.
DR PROSITE: PS50195; PX; 1.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain.
KW SH3 domain.
SQ SEQUENCE 349 AA; 38827 MW; 500F1E1247454101 CRC64;

Query Match 62.5%; Score 1219.5; DB 2; Length 349;
Best Local Similarity 67.4%; Pred. No. 8.1e-73;
Matches 242; Conservative 31; Mismatches 69; Indels 17; Gaps 4;

QY 1 MASPRPVSVQGAALVQIKLQTFATFSVRVSDSDTFFVRSSWDEFRQLKTKETFPVEA 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MASPRPVSAVAHALVQMDLQTFATFSVCSVSDSDTFFVRSSWDEFRQLKTKETFPVEA 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GLRRSDVLPKLLDAPLIGRGVTSGLRLQLETSYRLLATAERVARSPITITGFFA 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GLRRSEQVLPKLPDAPLLTRGHTGRLVRLDITVQALLATSEHILRSSALHGFFV 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 PQLDLPEALPPGSRVILPTPEEQPLSRAAGRLSIHSLEAQSRLCLQPFCTQDTRDPFQ 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 PKPLDLEPMLPPGSLVILPTPEE-PLSQPRGSLDIHSLEAQSIPCVQPFHTLDIRDPFH 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 AQAQESLDVLRHPSGHWLVENEDRQTAWFPAPYLEAAPGQREGGFSLGSGGPQFCAS 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 TKAQETLDIILRHPSGHWLVENEDQVWFPAPYLEEATVCOQESGLALQGSGRQFCCT 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 RAYESSRADELSVPAGARVRLVTSRGMWLCRYGVRAGLLPAVILRPEGLALLSGTGF 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 240 QAYEGSRDSELSVPSGARVRLVTSRGMWLCRYNGRTGLLPAMSLQPEGLGSLGRPGF 299
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 --RGGDDPAGEARGFPPEPQATAPPPTVTRPSFGAQSRCCCTWTRRAL-----ERRPR 352
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 300 PDSAGADKVAEDTI-----PPVPTPCMSALQSRCCCTTRALQEQGQGVTRPR 349
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
Q76EW0 PRELIMINARY; PRT; 253 AA.
AC Q76EW0;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE NADPH oxidase organizer 1 (Fragment).
GN Names=NOX01;
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley;
RX PubMed=14978110;
RA Kawahara T., Kuwano Y., Teshima-Kondo S., Takeya R., Sumimoto H.,
RA Kishi K., Tsunawaki S., Hirayama T., Rokutan K.;
RT "Role of nicotinamide adenine dinucleotide phosphate oxidase 1 in
RT oxidative burst response to toll-like receptor 5 signaling in large
RT intestinal epithelial cells";
RL J. Immunol. 172:3051-3058 (2004).
CC -1- SIMILARITY: Contains 1 SH3 domain.
DR EMBL: AB105906; BAD11766.1; -.
DR GO: GO:007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR001683; PX.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00787; PX; 1.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00326; SH3; 2.
DR PROSITE: PS50195; PX; 1.
DR PROSITE: PS50002; SH3; 1.
KW SH3 domain.
FT NON_TER 1 1
FT 253 253

SQ SEQUENCE 253 AA; 28536 MW; 5616C7C89BE73124 CRC64;

Query Match 47.5%; Score 927; DB 2; Length 253;
Best Local Similarity 71.8%; Pred. No. 1.2e-53;
Matches 186; Conservative 18; Mismatches 49; Indels 6; Gaps 3;

QY 30 WSDGSDTFVRRSWDEFRQLKTKETFPVEAGLLRRSDRVLPKLLDAPLIGRVGRTSRGL 89
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 WSDGSDTFVCRSWDEFRQLKTKENFPVEAGLLRRSDRLPKLPDAPLLSRGGRTGRGL 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 90 ARLOLETYSRRLLATAERVARSPITITGFPAPQPLDLEPALPGSRVILPTPEEQPLSRA 149
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 VRLLNTYVQELALANVQVSWSPVLTSFPEPRNQDLESALPSSVMVILPVP-EKPSRP 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 150 AGRLSIHSLEAQAQLRCLQPFCTQDTRDRPFAQAQESLDVLRHPSGHWLVENEDRQTAW 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 VNSLDIHSLEIQSLCCVHPFHTQDTRGFRHVQAQESLDVLRHPSGHWLVENEGQKAW 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 210 FPAPYLEEAPGQREGGFSLGSGGPQFCASRAYESSRADELSVPAGARVRLVLETSRGG 269
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 FPAPYL-EVAP-----EVGLMLQNSGTVCASRAYKGSRADELSVPAGARVRLVLETSRGG 234
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 WLCRYGDRAGLLPAVILRP 288
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 235 WLCRYCDQDGLLPVLLQP 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
Q9D747 PRELIMINARY; PRT; 239 AA.
AC Q9D747;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310034C04 product:2310034C04RIK PROTEIN (SNX28)
DE homolog (SNX28).
GN Name=Noxol; Synonym=Snx28;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
```


ID NCF1_HUMAN STANDARD; PRT; 390 AA.
 AC P14598; O43842; Q9BU90; Q9BX17; Q9BX18; Q9UDV9; Q9UMU2;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
 DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (NCF-47K) (47 kDa
 DE autosomal chronic granulomatous disease protein) (NOXO2).
 GN Name=NCF1;
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89332501; PubMed=2547247;
 RA Lomax K.J., Leto T.L., Nunoi H., Gallin J.I., Malech H.L.;
 RT "Recombinant 47-kilodalton cytosol factor restores NADPH oxidase in
 RT chronic granulomatous disease.";
 RL Science 245:409-412(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89386707; PubMed=2550933;
 RA Volpp B.D., Nauseef W.M., Clark R.A.;
 RT "Cloning of the cDNA and functional expression of the 47-kilodalton
 RT cytosolic component of human neutrophil respiratory burst oxidase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7195-7199(1989).
 RN [3]
 RP REVISIONS.
 RA Volpp B.D., Nauseef W.M., Clark R.A.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9563-9563(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90377229; PubMed=2398896;
 RA Rodaway A.R.F., Teahan C.G., Casimir C.M., Segal A.W., Bentley D.L.;
 RT "Characterization of the 47-kilodalton autosomal chronic granulomatous
 RT disease protein: tissue-specific expression and transcriptional
 RT control by retinoic acid.";
 RL Mol. Cell. Biol. 10:5388-5396(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97474758; PubMed=9329953;
 RA Gorlach A., Lee P.L., Roessler J., Hopkins P.J., Christensen B.,
 RA Green E.D., Chanock S.J., Curnutte J.T.;
 RT "A p47-phox pseudogene carries the most common mutation causing p47-
 RT phox-deficient chronic granulomatous disease.";
 RL J. Clin. Invest. 100:1907-1918(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20238075; PubMed=10772875; DOI=10.1006/bcmd.2000.0274;
 RA Chanock S.J., Roessler J., Zhan S., Hopkins P., Lee P., Barrett D.T.,
 RA Christensen B.L., Curnutte J.T., Goerlach A.;
 RT "Genomic structure of the human p47-phox (NCF1) gene.";
 RL Blood Cells Mol. Dis. 26:37-46(2000).
 RN [7]
 RP SEQUENCE FROM N.A., AND VARIANT GLU-259.
 RC TISSUE=Umbilical vein;
 RX MEDLINE=21606173; PubMed=11740866; DOI=10.1006/excr.2001.5404;
 RA Gu Y., Xu Y.C., Wu R.P., Souza R.F., Nwariaku F.E., Terada L.S.;
 RT "TNFalpha activates c-jun amino terminal kinase through p47(phox).";
 RL Exp. Cell Res. 272:62-74(2002).
 RN [8]
 RP SEQUENCE FROM N.A., AND VARIANTS GLY-99 AND ASN-166.
 RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
 RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
 RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
 RA Wyllie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
 RA Fowell G.A., Delehaanty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
 RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
 RA Ozeraky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
 RA Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
 RA Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Stromatt C.,
 RA Latraille P., Miller N., Johnson D., Murray J., Woessner J.P.,
 RA Wendt M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J.,
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
 RA Clifton S.W., Chisoe S.L., Marra M.A., Raymond C., Haugen E.,
 RA Gillitt W., Levy R., James R., Phelps K., Iadamoto S., Bubbs K.,
 RA Simms E., Zhou Y., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
 RA Waterston R.H., Wilson R.K.;
 RT "The DNA sequence of human chromosome 7.";
 RL Nature 424:157-164(2003).
 RN [9]
 RP SEQUENCE FROM N.A., AND VARIANTS GLY-99 AND ASN-166.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [10]
 RP SEQUENCE OF 8-31 FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96090243; PubMed=7581362;
 RA Harshman K., Bell R., Rosenthal J., Katcher H., Miki Y., Swenson J.,
 RA Chohami Z., Frye C., Ding W., Dayananth P., Eddington K., Norris F.H.,
 RA Ristow P.K., Phelps R., Hattier T., Stone S., Shaffer D., Bayer S.,
 RA Bussey C., Tran T., Lai M., Rostock P.R. Jr., Skolnick M.H.,
 RA Shattuck-Eidens D., Kamb A.;
 RT "Comparison of the positional cloning methods used to isolate the
 RT BRCA1 gene.";
 RL Hum. Mol. Genet. 4:1259-1265(1995).
 RN [11]
 RP STRUCTURE BY NMR OF 1-128.
 RX MEDLINE=21267481; PubMed=11373621; DOI=10.1038/88591;
 RA Hiroaki H., Ago T., Ito T., Sumimoto H., Kohda D.;
 RT "Solution structure of the PX domain, a target of the SH3 domain.";
 RL Nat. Struct. Biol. 8:526-530(2001).
 RN [12]
 RP VARIANT CGD GLN-42, AND VARIANT SER-262.
 RX MEDLINE=20575420; PubMed=11133775; DOI=10.1182/blood.V97.1.305;
 RA Noack D., Rae J., Cross A.R., Ellis B.A., Newburger P.E.,
 RA Curnutte J.T., Heyworth P.G.;
 RT "Autosomal recessive chronic granulomatous disease caused by defects
 RT arising in the NCF1 pseudogenes.";
 RL Blood 97:305-311(2001).
 CC -!- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are
 CC required for activation of the latent NADPH oxidase (necessary for
 CC superoxide production).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DISEASE: Defects in NCF1 are the cause of autosomal cytochrome-b-
 CC positive chronic granulomatous disease type 1 (CGD) [MIM:233700];
 CC an autosomal recessive form.
 CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.


```
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF0018; SH3.1; 2.
DR PRINTS; PR00498; P47PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD00066; SH3; 2.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS0002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 389 AA; 44679 MW; 5296142988292F03 CRC64;

Query Match 15.7%; Score 307; DB 2; Length 389;
Best Local Similarity 27.1%; Pred. No. 2e-12;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVMSDGSDFVRRSWDEFRQLKTLKTFVVEAGLLRRSDRVLPKLLDAPLLGRVG 83
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 YMFVVKQDLSEKVVYRKFTIEYFHKMLKEMFPIEAGEIHTENRVIPH-LPAPRWYDQG 84
QY 84 RTSRGLARLQLETYSRRLLATAERVARSPITITGFFAPQPLDLEPALPGSRVILPTPEE 143
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 RAAE--SRQGLTEYFNSLMGLPVKISRCPHLLNFFKVRPDDLK--LPNDSQVKKP---E 137
QY 144 QPLSRAAGRLSIHSLA----QSLRCLQPFCTQDRDRPFQAQAESLDVLLRHPSGWGL 199
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 TYLTAKDGKKNVADITGPILQTYRAIADY--EKGSKTMTVATGDDVVVKSSESGWVF 195
QY 200 VENEDRQTAWFPAPYLEEA-APCGREGGSLSSGPGQFCASRAYESSRADELSPAGAR 258
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 196 CQMKTKR-GWVPASYLEPLSDPEAEDPDNY--AGEPYVTKAYAAVEDESLSEGEA 252
QY 259 VRVLETSRGMWLCRYGDRAGLLPALLVPEGLGALLSGTGFRGGDDPAGAR----- 311
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 IEVIHKLDDGWWVRKGDITGYFPSPMYLKAGEEITQAQRQIRGRCGAPRRSIRNAQSI 312
QY 312 -----GFEPSQA-----TAPPPTVTPRPSG 333
Db 313 HORSRKLSDQTYRRNSVRFLOQRPRRPGPQSGTGTGKDNPTPRVKQPAVPPRPSSD 372
QY 334 AIQSRCCVTTRAL 347
Db 373 LILHRCSTESTKRKL 386

RESULT 12
Q7TMH0 PRELIMINARY; PRT; 390 AA.
AC Q7TMH0
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Neutrophil cytosolic factor 1.
GN Name=Ncf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```



```

ID Q99M65 PRELIMINARY; PRT; 389 AA.
AC Q99M65;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE P47 phox (Neutrophil cytosolic factor 1).
GN Name=Ncf1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Tanabe M., Radmark O.P.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RX MEDLINE=22397839; PubMed=12461526; DOI=10.1038/ng1058;
RA Olofsson P., Holmberg J., Tordsson J., Lu S., Akerstrom B.,
RA Holmdahl R.;
RT "Positional identification of Ncf1 as a gene that regulates arthritis
RT severity in rats."
CC -!- SIMILARITY: Contains 2 SH3 domains.
DR EMBL; AY029167; AAK31797.1; -.
DR EMBL; AF547392; AAO32680.1; -.
DR HSSP; P14598; 1UEC.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001655; P47PHOX.
DR InterPro; IPR001683; PX.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF00018; SH3; 1; 2.
DR PRINTS; PR00498; P47PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 2.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS0195; PX; 1.
DR PROSITE; PS00002; SH3; 2.
KW SH3 domain.
SQ SEQUENCE 389 AA; 44742 MW; BFC75842E53E68A4 CRC64;

Query Match
Best Local Similarity 27.1%; Score 306; DB 2; Length 389;
Matches 101; Conservative 55; Mismatches 155; Indels 62; Gaps 14;

QY 24 FAFSVRWSGSDTFVRRSWDEFRQLKTKLKEFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83
D 26 YMFVVKQDLSKVVYKFTTEIYEPFKMLKEMFPIEAGEIHTENRVIHP-LPAPRWYDQG 84
QY 84 RTSRGLARLQLETSYRLLATAERVARSPITGFAFPQDLEPALPGSRVILPTPEE 143
D 85 RAAE--SRQGLTTEYFNLSMLGLPMKISRCPHLLNFFKVRPDDLK--LPNDSQVKKP--- 137
QY 144 QPLSRAAGRLSIHSLA-----QSLRCLQFCTQTRDRPFQAQAESLDVLLRHPSGWWL 199
D 138 TYLTAKDGKNNVADIMGPIILQTYRAIDY--EKSKTEMTVATGDVVDVVEKSSGWWF 195
QY 200 VENEDROTAWFPAPYLEA-AFGQREGGSPISGGPQFCASRAYESSRADELSVPAGAR 258
D 196 CQMKTKR-GWVPASYLEPLDSPEADEPDPNY--AGEPVVTKAYAAVEDEVSLSGEA 252
QY 259 VRVLETSRGRWLCRYGDRAGLLPAVLRPEG-----LGA----- 293
D 253 IEVIHKLDGWWVRKGDITGYFPFMYLQKAGEITQARQIRSGAPRRSTINNAOSI 312
QY 294 -----LLSGTGFRCGD-----DPAGEARGFPE-----PSQATA-PPPTVPTRPSPCA 334
D 313 HQRSKRSLSDTYRRNSVRFLOORRPARPGQSPDSKDNFSTTPRAKQPAPVPPRSSDL 372
```

```

QY 335 IQSRCCCTVTRRAL 347
D 373 ILHRCSTSTRKL 385

RESULT 14
NCFL MOUSE
ID NCFL_MOUSE STANDARD; PRT; 390 AA.
AC Q03014; Q070144; Q9J134;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (NCF-47K).
GN Name=Ncf1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX MEDLINE=94164697; PubMed=8119734;
RA Jackson S.H., Malech H.L., Kozak C.A., Lomax K.J., Gallin J.I.,
RA Holland S.M.;
RT "Cloning and functional expression of the mouse homologue of
RT p47phox."
RN Immunogenetics 39:272-275 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=98149672; PubMed=9490028;
RA Mizuki K., Kadamatsu K., Hata K., Ito T., Fan Q.-W., Kage Y.,
RA Fukumaki Y., Sakaki Y., Takeshige K., Sumimoto H.;
RT "Functional modules and expression of mouse p40(phox) and p67(phox),
RT SH3-domain-containing proteins involved in the phagocyte NADPH oxidase
RT complex."
RN Eur. J. Biochem. 251:573-582 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Green E.D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NCF2, NCF1, and a membrane bound cytochrome b558 are
CC required for activation of the latent NADPH oxidase (necessary for
CC superoxide production).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
CC -!- SIMILARITY: Contains 2 SH3 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11455; AAA50469.1; -.
DR EMBL; AB002663; BAA25649.1; -.
DR EMBL; AF267747; AAF90134.1; -.
DR HSSP; P14598; 1UEC.
DR MGD; MGI:97283; Ncf1.
DR GO; GO:0016175; F:superoxide-generating NADPH oxidase activity; IEA.
DR GO; GO:0008283; P:cell proliferation; IMP.
DR GO; GO:0016066; P:cellular defense response (sensu Vertebrata); IMP.
DR GO; GO:0006954; P:inflammatory response; IMP.
DR GO; GO:0006691; P:leukotriene metabolism; IMP.
DR GO; GO:0009617; P:response to bacteria; IMP.
DR InterPro; IPR001655; P47PHOX.
DR InterPro; IPR001683; PX.
DR Pfam; PF00787; PX; 1.
```

```
DR Pfam; PF00018; SH3; 2.  
DR PRINTS; PR00498; P47PHOX.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR ProDom; PD000066; SH3; 2.  
DR SMART; SM00312; PX; 1.  
DR SMART; SM00326; SH3; 2.  
DR PROSITE; PS0195; PX; 1.  
DR PROSITE; PS0002; SH3; 2.  
KW Repeat; SH3 domain.  
FT DOMAIN 4 125  
FT DOMAIN 211 254 Asp/Glu-rich (highly acidic).  
FT DOMAIN 292 390 Arg/Lys-rich (highly basic).  
FT DOMAIN 162 215 SH3 1.  
FT DOMAIN 226 285 SH3 2.  
FT CONFLICT 161 161 Y -> H (in Ref. 1).  
FT CONFLICT 343 345 GOL -> RAA (in Ref. 1).  
FT CONFLICT 344 344 Q -> P (in Ref. 3).  
SQ SEQUENCE 390 AA; 44698 MW; A383DB953839CFBC CRC64;  
  
Query Match 15.6%; Score 303.5; DB 1; Length 390;  
Best Local Similarity 25.7%; Pred. No. 3.4e-12;  
Matches 96; Conservative 55; Mismatches 160; Indels 63; Gaps 11;  
  
QY 24 FAFSVWSDGSDTFVRRSWDEFRQLKTKETFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83  
DB 26 YMFVWKQDLSEKVIYRRFTETIYEFHMLKEMFPPIEAGEIHTENRVIPIH-LPAPRWFDGQ 84  
  
QY 84 RTSRGLARQLLETYSRRLLATAERVARSPITITGFAPOPLDLEPALPGSRVILPTPEE 143  
DB 85 RAAE--SRQGLTEYNGLMGLPVKISRCPLHLLDFKVRPDDLK--LPTDSQAKK--E 137  
  
QY 144 QPLSRAAGRLSTHSLA-----QSLRCLQPFCTQDTRDPFQAQAQBSLDVLLRHPSGWWL 199  
DB 138 TVLVPKDGNVADITGPILQTYRAIADY--EKSGTGMTVATGDVVDVVEKSESGWVF 195  
  
QY 200 VENEDQTAWFAPYLEEA-APCQREGGSPSLGSSGPQCASPAYESSRADELSPAGAR 258  
DB 196 CQMKTKR-GWVPASYLEPLDSDPEAEDDPFNY--AGEPYVTKIYAAVBEDEMSLSEGEA 252  
  
QY 259 VRVLETSRGMWLCRYGDRAGLLPALLRPEGLGALLSGTGPRGGDDPAGEA----- 310  
DB 253 IEVIHKLDDGWWVRKEDITGTFPSYLOKAGEITQAORQIKRGAPRRSIRNAHQSI 312  
  
QY 311 --RGFPEPSQAT-----APPTVTPRPSG 333  
DB 313 HQRSKRRLSQDTYRRNSVRFLOQRRRPGRGQLSTDGTDKNPSTPRVKPQPAVPPRPSD 372  
  
QY 334 ATQSRCTVTTRAL 347  
DB 373 LILHRCSTETKRL 386  
  
RESULT 15  
Q95MNO PRELIMINARY; PRT; 391 AA.  
AC Q95MNO;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE NADPH oxidase p47-phox.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21676683; PubMed=11818454;  
RA Gauss K.A., Mascolo P.L., Siemsen D.W., Nelson L.K., Bunger P.L.,  
RA Pagano P.J., Quinn M.T.;  
RT "Cloning and sequencing of rabbit leukocyte NADPH oxidase genes  
RT reveals a unique p67(phox) homolog."  
RC J. Leukoc. Biol. 71:319-328(2002).  
CC -!- SIMILARITY: Contains 2 SH3 domains.
```

```
DR EMBL; AF324409; AAK58851.1; -.  
DR HSSP; P14598; IUEC.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR InterPro; IPR001655; P47PHOX.  
DR InterPro; IPR001683; PX.  
DR InterPro; IPR001452; SH3.  
DR Pfam; PF00787; PX; 1.  
DR Pfam; PF00018; SH3 1; 2.  
DR PRINTS; PR00498; P47PHOX.  
DR ProDom; PD000066; SH3; 2.  
DR SMART; SM00312; PX; 1.  
DR SMART; SM00326; SH3; 2.  
DR PROSITE; PS0195; PX; 1.  
DR PROSITE; PS0002; SH3; 2.  
KW SH3 domain.  
SQ SEQUENCE 391 AA; 44907 MW; 22DDA523BC871CCA CRC64;  
  
Query Match 15.5%; Score 302; DB 2; Length 391;  
Best Local Similarity 26.1%; Pred. No. 4.3e-12;  
Matches 98; Conservative 54; Mismatches 158; Indels 66; Gaps 12;  
  
QY 24 FAFSVWSDGSDTFVRRSWDEFRQLKTKETFPVEAGLLRRSDRVLPKLLDAPLLGRVG 83  
DB 26 YMFVWKQDLSEKVIYRRFTETIYEFHMLKEMFPPIEAGEIHTENRVIPIH-LPAPRWFDGQ 84  
  
QY 84 RTSRGLARQLLETYSRRLLATAERVARSPITITGFAPOPLDLEPALPGSRVILPTPEE 143  
DB 85 RAAE--SRQGLTEYCGALMALPVKISRCPLHLLDFKVRPDDLK--LPADSOA--KKPET 138  
  
QY 144 QPLSRAAGRLSTHSLAQAQSLRCLQPFCTQDTRDPFQAQAQ--ESLDVLLRHPSGWWL 202  
DB 139 YLMSKDAKSNVADITGPILQTYRAIANYEKHSPSEMAATGDVVDVVEKSESGWVF 198  
  
QY 203 EDROTAWFPAPYLEEA-APCQREGGSPSLGSSGPQCASPAYESSRADELSPAGARV 261  
DB 199 KTKR-GWVPASYLEPLDSDPEAEDDPFNY--AGEPYVTKIYATAVEDEBVSQAQGEAIEV 255  
  
QY 262 LETSRGMWLCRYGDRAGLLPALLRPEG-----LGA----- 293  
DB 256 IHKLLDGGWWVRKEDVTGYFSPMFLQKSQDLTQAORQIKRGAPRRSIRNAHQSI 315  
  
QY 294 -----LLSGTGPRGGDDPAGEARGPPEPSQATAPPTVTPRPS 331  
DB 316 ARKRLSQDTYRRNSVRYLQRRRLPPSQRGASPLQEE---PRP-QSTKQPAPVPPRPS 371  
  
QY 332 PGAIQSRCCTVTTRAL 347  
DB 372 ADLILNRCSSTKRL 387  
  
Search completed: May 28, 2005, 07:44:16  
Job time : 173 secs
```

This Page Blank (uspto)